



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 107949

TO: Terra Gibbs
Location: cm1/12a12/11e12
Art Unit: 1635
Sunday, November 16, 2003

Case Serial Number: 10/005337

From: Toby Port
Location: Biotech-Chem Library
CM1-6A04
Phone: 308-3534

toby.port@uspto.gov

Search Notes

Dear Examiner Gibbs,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

THIS PAGE BLANK (USPTO)



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 - Circ. Desk



THIS PAGE BLANK (USF-14)

STIC-Biotech/ChemLib

107 949

From: Gibbs, Terra
Sent: Sunday, November 09, 2003 2:54 PM
To: STIC-Biotech/ChemLib
Subject: Sequence search request...

RECEIVED

NOV 10 2003

Could you please do a regular search of SEQ ID NOs: 1 and 2 of USSN 10/005337?

(STIC)

Also could you please search do an oligomer search SEQ ID NOs: 1 and 2, length limited to 50 nucleobases or less.

Thank You.

Terra Cotta Gibbs, Ph.D.
Art Unit 1635
CM1, 12A12
703-306-3221

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (see text)



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 – Circ. Desk



THIS PAGE BLANK (USPTO)

Schreiber, David

108482

From: Gibbs, Terra
Sent: Sunday, November 09, 2003 2:52 PM
To: Schreiber, David
Subject: Sequence search request...

Hi David, I have a sequence search request.

NO CRF

The request is for % identity of SEQ ID NOs: 1 and 2 of USSN 10/005337 to SEQ ID NO:3 of PCT US99/20730.

I need SEQ ID NO:1 to be at least 93% identical with SEQ ID NO:3.
I also need SEQ ID NO:2 to be at least 80% identical with SEQ ID NO:3.

NO 601099,900

Is this possible?
Thank You!

Terra Cotta Gibbs, Ph.D.
Art Unit 1635
CM1, 12A12
703-306-3221

run seq 1 & 2

THIS PAGE BLANK (user)

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

STAFF USE ONLY		Type of Search	Vendors and cost where applicable
Searcher: <u>R. Schreiber</u>	NA Sequence (#) <u>2</u>	STN _____	
Searcher Phone #: <u>308-4292</u>	AA Sequence (#) _____	Dialog _____	
Searcher Location: <u>CMI 6A03</u>	Structure (#) _____	Questel Orbit _____	
Date Searcher Picked Up: _____	Bibliographic _____	Dr. Link _____	
Date Completed: <u>11/18</u>	Litigation _____	Lexis-Nexis <u>CompuGen GCG</u>	
Searcher Prep & Review Time: <u>14</u>	Fulltext _____	Sequence Systems _____	
Client Prep Time _____	Patent Family _____	WWW Internet _____	
On-line Time: <u>54</u>	Other _____	Other (specify) _____	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 03:46:27 ; Search time 4312.25 Seconds
(without alignments)
11689.358 Million cell updates/sec

Title: US-10-005-337A-2

Perfect score: 2074

Sequence: 1 ctgcagcaagttacttaatg.....acaagactcttcagccaac 2074

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST.*
1: em_estba.*
2: em_estba.*
3: em_estin.*
4: em_estin.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_estc.*
9: gb_est1.*
10: gb_est2.*
11: gb_estc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_man.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_ptg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	461.2	22.2	522	28	AQ480395
2	260.6	12.6	586	28	AZ295534
3	241	11.6	735	9	AU139209
4	181	8.7	371	13	BY091041

5	158.4	7.6	353	13	BY079997
6	98.6	4.8	343	28	AQ584530
7	98.6	4.8	827	14	CD244827
8	95.8	4.6	502	28	AQ207760
9	95	4.6	412	9	AA494075
10	94.2	4.5	464	9	AA917468
11	94.2	4.5	2772	11	BC038630
12	93.8	4.5	315	9	AI039619
13	93.6	4.5	369	10	BF738217
14	93.6	4.5	464	10	BF829139
15	93.6	4.5	723	28	AZ335882
16	93.2	4.5	488	10	BF415552
17	93.2	4.5	532	29	BZ460574
18	92.8	4.5	477	28	AZ406305
19	92.6	4.5	380	9	AA947380
20	92.4	4.5	534	10	BF724783
21	92.2	4.4	387	10	BF400140
22	92.2	4.4	453	28	AQ187593
23	92	4.4	563	12	BI438856
24	92	4.4	567	12	BI439182
25	92	4.4	950	10	BG390793
26	91.8	4.4	451	28	AQ108862
27	91.6	4.4	419	9	AA468422
28	91.6	4.4	717	28	AQ343576
29	91.4	4.4	442	28	AQ219900
30	91.4	4.4	662	28	AZ365914
31	91.4	4.4	2073	11	AF289601
32	91.2	4.4	470	9	AI814961
33	91	4.4	357	9	AW002875
34	91	4.4	464	9	AA845825
35	91	4.4	556	29	BZ606409
36	91	4.4	587	9	AV720761
37	91	4.4	587	9	AV720765
38	91	4.4	589	9	AV720514
39	91	4.4	639	28	AQ109988
40	91	4.4	988	10	BF791763
41	91	4.4	1017	12	BM472183
42	90.8	4.4	284	14	T39841
43	90.8	4.4	541	28	AQ895064
44	90.8	4.4	821	14	CD520493
45	90.6	4.4	430	9	AI288531

ALIGNMENTS

RESULT 1
AQ480395/c

LOCUS: AQ480395 522 bp DNA linear GSS 23-APR-1999
DEFINITION: RPCI-11-236B22.TV RPCI-11 Homo sapiens genomic clone RPCI-11-236B22
ACCESSION: AQ480395
VERSION: AQ480395.1 GI:4662514
KEYWORDS: GSS.
SOURCE: Homo sapiens (human)
ORGANISM: Homo sapiens

REFERENCE: 1 (bases 1 to 522)
AUTHORS: Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.

TITLE: Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

JOURNAL: Map Building

COMMENT: Unpublished
Other_GSSs: RPCI-11-236B22.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.

FEATURES

Location/Qualifiers

1..522
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7590285"
/db_xref="taxon:9606"
/clone="RPCI-11-236B22"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI; RPCI11 Human Male BAC Library"

168 a 100 c 113 g 141 t

Query Match 22.2%; Score 461.2; DB 28; Length 522;

Best Local Similarity 97.7%; Pred. No. 2.5e-19;

Matches 510; Conservative 1; Mismatches 5; Indels 6; Gaps 4;

QY 971 GTTACCATTAATCAATC-AGATTCCTCGGAGTGGGC-CAGGATCTGTATTTCTGAC 1028
|||||
DB 522 GTTACCATTAATCAATCAGAAATTCCTCGGAGTGGGCACAGGGATCTGTATTTCTGAC 463
|||||
QY 1029 AAGCTCCACAGGTGATTCCTTTCCACAGCATTTGAGAACTTCAGCTCAATGACCTAA 1088
|||||
DB 462 AAGCTCCACAGGTGATTCCTTTCCACAGCATTTGAGTACTTCAGCTCAATGACCTAA 403
|||||
QY 1089 TCAGAGTCTCGCATTCGTAATATCTGTCTCATTTTBTATATATATATATATATATTT 1148
|||||
DB 402 TCAGAGTCTCGCATTCGTAATATCTGTCTCATTTTBTATATATATATATATATATTT 343
|||||
QY 1149 GTGTAGAGATGGATTTTGCATTTGCCAGGTGTCAGCTAGTATTGAACTCTAGCTAAGCA 1208
|||||
DB 342 TTGTAGAGATGGATTTTGCATTTGCCAGGTGTCAGCTAGTATTGAACTCTAGCTAAGCA 283
|||||
QY 1209 ATCTCTCTCTCTGCTCCCAAAATGTTGGATTACAGGTGTAGCCACTGCACCCGGC 1268
|||||
DB 282 ATCTCTCTCTCTGCTCCCAAAATGTTGGATTACAGGTGTAGCCACTGCACCCGGC 223
|||||
QY 1269 TGATAGCTGGTTTCATTTACTCTTATTTCTTGACCACTCTGATCCATTTTGAAGTAAAT 1328
|||||
DB 222 TGATAGCTGGTTTCATTTACTCTTATTTCTTGACCACTCTGATCCATTTTGAAGTAAAT 163
|||||
QY 1329 GCTCCATTTATTTATGCTGTTTGTAGAACAGGTGAAGCATGTCATGTGCTA---ATGCCAG 1385
|||||
DB 162 GCTCCATTTATTTATGCTGTTTGTAGAACAGGTGAAGCATGTCATGTGCTAATTATGCCAG 103
|||||
QY 1386 TCACATCATAAAGAAAGTGATTTACTCAATGCTTTCAATGCTTTATATATGATGTAAG 1445
|||||
DB 102 TCACATCATAAAGAAAGTGATTTACTCAATGCTTTCAATGCTTTATATATGATGTAAG 43
|||||
QY 1446 GTGGCATGTCATGGGGCCCTATTAG-CCAGACATCACTCCA 1486
|||||
DB 42 GTGGCATGTCATGGGGCCCTATTAGCCCCAGACATCACTCCA 1

RESULT 2
AZ295534
LOCUS
DEFINITION
RPCI-23-105P3.TJ RPCI-23 Mus musculus genomic clone RPCI-23-105P3,
genomic survey sequence.
ACCESSION
AZ295534
VERSION
AZ295534.1 GI:9537319
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE

AUTHORS

Zhao,S., Niernan,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
and Fraser,C.M.

Mouse BAC End Sequences from Library RPCI-23

TITLE

JOURNAL

COMMENT

Other_GSSs: RPCI-23-105P3.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)

or from Resea ch Genetics (info@resgen.com). BAC end page:

http://www.tigr.org/cdb/bac_ends/mouse/bac_end_intro.html

Plate: 105 row: P column: 3

Seq primer: SP6

Class: BAC ends.

FEATURES

Location/Qualifiers

1..586
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-105P3"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site 1:
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACE3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies). "

BASE COUNT 141 a 148 c 154 g 143 t

ORIGIN

Query Match 12.6%; Score 260.6; DB 28; Length 586;

Best Local Similarity 75.9%; Pred. No. 1.8e-07;

Matches 403; Conservative 0; Mismatches 114; Indels 14; Gaps 6;

QY 1533 CTTCCCTCAGGCTTTTACCCAGGGAATAGGATGCC---TGGGACAAGTTTCCCTAA 1589
DB 21 CTTCCCTCAGGCTGATTATCCCAATAGGATGCCCAAGACAACATCTCCAGCCAA 80
QY 1590 GTGAAGTGTGATAAGTCTGCTTATCAGAAAGATATTACTGGGGGTGTGATGTAGGGC 1649
DB 81 CTGGAGTGTGATAAGTCCAGTTATCAGAAAGATATGGCTGTAAAGTGTGATGCACAGTGC 140
QY 1650 ATCTCATTTTCTGATAGGTAGTCATATGAAAGCTGACAAAGAA---AAAAAGGGCAGTG 1707
DB 141 ---TTGCATTTTCTTGATAGCTTAGTCATTATTATCTGACAAAGAGAAAGAGCAGCG 198
QY 1708 ATGTGGTGAATGTCAACAGACAGCTGCCCTGAC-TCTTACCAATAGGATGACTTGC 1766
DB 199 ATGTGGTGAATATTAAACAGGAGCTGTCCTCGCTTCCCATAGCTGGATGACTCGC 258
QY 1767 ATTGCTGACGATGTGATCACCACCAAGGAATGGCCCTCTCACATTTCTTCTGATTCA 1826
DB 259 ATTGCTGACGCTGGTGGTCACTGCCAAGGAATGACCTCTCACATTTCTTCTGATTGC 318
QY 1827 CATATTTCAGCAGGGTAGCTTCTCTCCCTTCCCTTTCAGCTTCCAGACACTGAGTCT 1886
DB 319 CATACGCCCGCGC-----CAGCTTGTCTATCTCCCTTGGGCTTCCAGACACTAAGTCT 373

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1. 371
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="K630084E21"
/tissue_type="heart"
/dev_stage="10 days neonate"
/clone_lib="RIKEN full-length enriched, 10 days neonate heart"

BASE COUNT 91 a 94 c 105 g 81 t

ORIGIN

Query Match 8.7%; Score 181; DB 13; Length 371;
Best Local Similarity 81.3%; Pred. No. 0.012; Mismatches 50; Indels 7; Gaps 3;
Matches 248; Conservative 0; TGGGATGACTCGCAATTCGAGCGGTGTCTCACTCGGAAGGATGACCCCTCTCACATT 64
1754 TAGGATGACTTCGATTCGATTCGAGCGATGTGATCACCACCAAGGAATGCGCTCTCACATT 1813
Db 5 TGGGATGACTCGCAATTCGAGCGGTGTCTCACTCGGAAGGATGACCCCTCTCACATT 64
1814 TCTTCTGATTCACATATTCAGAGGTTAGTTGTCTCCCTCTTCCAGCTTCC 1873
Db 65 TCTTCTGATTCACATATTCAGAGGTTAGTTGTCTCCCTCTTCCAGCTTCC 119
1874 ACAGACTGATTCGAAATTCACCTGCTCTGAGTTCGCTCCCTTAATGGCGGCGG 1933
Db 120 ACAGACTGATTCGAAATTCACCTGCTCTGAGTTCGCTCCCTTAATGGCGGCGG 179
1934 AGTGTACTTCGTTCCCAAGTTGAAGATTATCTCACCGGCCCTATATAGCTG 1993
Db 180 GGTGTGACTTGGCTTCCAGGCTGGAAGATTATCTCACCGGCCCTATATAGCTG 238
1994 ACCGGTGTGGAGGGCCAGCGGCAACTCTCAGGATTCCTTC-CAGCAGAGAAAC 2052
Db 239 GCTGGTGTGGAGGGCTCCACAGGCGGCTCCAGGGGTTTCATCCACAGAGAGAAAC 298
2053 ATACA 2057
Db 299 ATAGA 303

RESULT 5

BY079997

LOCUS

DEFINITION

BY079997 RIKEN full-length enriched, 10 days neonate heart Mus musculus cDNA clone K630022E19 5', mRNA sequence.

ACCESSION

BY079997

VERSION

BY079997.1

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

REFERENCE

AUTHORS

1. (bases 1 to 353)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tonari, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest

A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKensie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Vezardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yamagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sakai, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp
URL: <http://genome.gsc.riken.go.jp/>

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sakai, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1. 353
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="K630022E19"
/tissue_type="heart"
/dev_stage="10 days neonate"
/clone_lib="RIKEN full-length enriched, 10 days neonate heart"

BASE COUNT 82 a 88 c 102 g 79 t 2 others

ORIGIN

Query Match

Best Local Similarity

7.6%; Score 158.4; DB 13; Length 353;

73.9%; Pred. No. 0.27;

Db 582 AATAACTAAGCACGGTGTCTCGCTTCGGCAGCACATATACTAAAATTGGAACGATACAGAG 523

Qy 318 AAGATTAGCATGGACTCTGTGCAAGAAATGACACAAAATTTGTGAAAATTTCATATATT 377

Db 522 AAGATTAGCATGGCCCTTCGGCAGGATGACACGAAATTTCTGTGAAGCGTTCCATATTATTA 463

Qy	378	AAAAATAAATAAATAAAGAGAAAAAGGAAAAAATTAAAAAGAAAAATA	426
Db	462	AAAAAATAAAAAAAGAAAAAATAAATCCACAAAAAACAACAAAAACAAGA	414

RESULT 8	502 bp	DNA	linear	GSS 18-SEP-1998
AQ207760				
LOCUS	AQ207760			
DEFINITION	HS_3026_B1_G10_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3026 Col=19 Row=N, genomic survey sequence.			

BASE COUNT	130 a	117 c	110 g	140 t	5 others
ORIGIN					
Query Match	4.6%	Score 95.8	DB 28	Length 502	
Best Local Similarity	71.9%	Pred. No. 1.1e+03			
Matches 138; Conservative 0; Mismatches 53; Indels 1; Gaps 1;					
Qy	1135	TATATATAGTATTTTGCGTAGAGATGGATTTTGCCTATGTTGCCATGTTGCCACGCTAGTATTTCGAAC	1194		
Db	291	TAATTTTGTATTTTTTGTAGACATGGTTTTTGCCTATGTTGCCATGTTGCCACGCTGGCTTCGGAAC	350		
Qy	1195	TCCTAAGCT-AAGCAATCTTCCTGTCTCTGCCTCCCAAAATGTTGGGATTTACAGGTGTAA	1253		
Db	351	TCCTGAGCTCAAGCAGTCTGCCTGCCTCAGCCTCGCAAGTGTGGGATTTACAGGCATAA	410		
Qy	1254	GCCTAGCACCCGGCTCATAGCTGGTTTCATTTACTCTATTTCTTGACCACTCTGTATCCA	1313		
Db	411	GCCTATGCACCCAGCAGGAGAAAGATNTAAAACTATCTTTCTGTGAAGAAAAAAGTACA	470		

[illegible]

BASE COUNT	118 a	82 c	114 g	98 t
ORIGIN				
Query Match	4.6%	Score 95	DB 9	Length 412
Best Local Similarity	74.8%	Pred. No. 1.4e+03		
Matches 119; Conservative	0	Mismatches 40	Indels 0	Gaps 0
Qy 1129	CATATATATATATAGTATTTTGTGTGAGAGATGGGATTTTGCATGTTGCCAGGCTAGTA	1188		
Db 272	CCTAGCTTATTTTGTATTTTGTAGAGATGGGTTTTCGCAATGTTTACCCAGGTTGATC	213		
Qy 1189	TTGAACTCCTAAGCTAAGCAATCTTCTGTCTGTGCTCCCAAAATGTTGGGATTACAGG	1248		
Db 212	TCAAACCTCTGAGCTCAGCAGTCTGCGCTGCTCGGCCCTCCCAAAGTCTGGGATTACAGG	153		
Qy 1249	TGTAAGCCACTGCACCCGGCTGATAGCTGGTTTTCATTTA	1287		
Db 152	CGTGAGCCACACACCTGGGCCAATAACTTGCAAATTTTA	114		

RESULT 10				
AA917468				
LOCUS	AA917468	464 bp	linear	EST 10-JUN-1998

DEFINITION ol51f11.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:1527021.3' similar to SW:TS6 HUMAN P98066 TUMOR NECROSIS
FACTOR-INDUCIBLE PROTEIN TSG-6 PRECURSOR ;contains Alu repetitive
element'; mRNA sequence.
AA917468
ACCESSION AA917468.1 GI:3057358
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 464)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 417 Std Error: 0.00
Seq primer: -40m13 fwd. RT from Amersham
High quality sequence stop: 345.
Location/Qualifiers
FEATURES
source
1..464
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1527021"
/lab_host="DH10B"
/clone_lib="Soares_NFL_T_GBC_S1"
/notes="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbH19W, testis NHT, and B-cell
NCI-CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. 682632-687239,
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 118 a 101 c 96 g 149 t
ORIGIN
Query Match 4.5%; Score 94.2; DB 9; Length 464;
Best Local Similarity 70.4%; Pred. No. 1.4e+03;
Matches 126; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 1117 TCTCATTTTBT CATATATATATAGTATTGTGGTAGAGATGGGATTTGGCATGTTG 1176
DB 1 TTTAAATTTTAAATTTTAAATTTTAAATTTTTCATAGAGATGAGGTCTTGCTATGTTG 60
QY 1177 CCAGGCTAGTATGAACCTCTAAGTAAGCAATCTTCCTGTCTCTGCCCTCCCAAAATGT 1236
DB 61 CCAGGCTTGTGTTTAACTCCTGAGCTCAAGAGATCTCTGCTCTGCTCTCCCAAGTGA 120
QY 1237 TGGATTACAGGTGTAAGCCACTGACCCGGCTGATGCTGGTTTCATTACTCTATT 1295
DB 121 TGGGATTACAGGAGTGAGCCACTGCGCGGGTGAATCTCTGGTTTATGACTTGCTTT 179
RESULT 11
BC038630
LOCUS
DEFINITION BC038630 2772 bp mRNA linear HTC 04-MAR-2003
Homo sapiens, similar to hypothetical protein FLJ20489, clone
IMAGE:5263792, mRNA.
ACCESSION BC038630
VERSION BC038630
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2772)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
Contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 73 Row: 1 Column: 4
This clone has the following problem: retained intron.
Location/Qualifiers
FEATURES
source
1..2772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5263792"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"
BASE COUNT 682 a 523 c 602 g 865 t
ORIGIN
Query Match 4.5%; Score 94.2; DB 11; Length 2772;
Best Local Similarity 68.9%; Pred. No. 4.8e+02;
Matches 144; Conservative 0; Mismatches 63; Indels 2; Gaps 1;
QY 1129 CATATATATATATAGTATTGTGGTAGAGATGGGATTTGCCATGTTGCCAGGCTAGTA 1188
DB 997 CACCTATTTTATTCGATTTTATAGACAGCGGTTTGGCCATGTTGCCAGGCTGGTC 1056
QY 1189 TTGAACCTCTAAGCTAAGCAATCTTCTGTCTCTGCTCCCAAAATGTTGGGATTACAGG 1248
DB 1057 TTCAACCTCTGACCTCAGGTGAACCTTCTGCTCGGCTCCCAAAAGTCTGGGATTACAGG 1116
QY 1249 TGTAGCCACTGCACCGGCTGATA--GCTGGTTTCATTACTCTATTTCTTGACCACCTC 1306
DB 1117 CGTAGGCCACACCGCCGCCCTATATTCAGTTTTAAATAACACCATCTGAAGGCCACCA 1176
QY 1307 TGATCCATTTTGAAGTAAATAATGCTCAA 1335
DB 1177 AGAGTCCAGTCCCGAGTAGGAATATTCAA 1205
RESULT 12
AI039619/c
LOCUS
DEFINITION AI039619 315 bp mRNA linear EST 30-JUN-1998
OX28f05.x1 Soares_total_fetus Nb2HF8_9w Homo sapiens cDNA clone
IMAGE:1657865 3' similar to WP:T20D3.3 CE03672 ;, mRNA sequence.
AI039619
ACCESSION AI039619.1 GI:3278813
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 315)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 296.

FEATURES

source
 1. 315
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1657665"
 /dev_stage="8-9 weeks"
 /lab_host="DH108"
 /clone_lib="Soares total fetus Nb2HF8 9w"
 /note="Vector: pT7T3D-PaC (Phatmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from pooled 8-9 week
 (total) fetus material with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTGAAGTGGAGCGCGCTTAATTTTTTTTTTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaudo."

BASE COUNT 61 a 65 c 69 g 120 t
 ORIGIN
 Query Match 4.5%; Score 93.8; DB 9; Length 315;
 Best Local Similarity 72.2%; Pred. No. 1.9e+03;
 Matches 122; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
 QY 258 AATAGTAGTACTCTGTTGCTTCAGCAGGACATATCAATAAATAGGAGCTATCAAG 317
 DB 177 ATAATAAGCAGCGTCTGCTTCGGCAGCAGCATATCAATAAATGGACATACAG 118
 QY 318 AAGATTAGCATGCTCTGTGCAAGATGACACAAAATTTGTGAACATTCATATATT 377
 DB 117 AAGATTAGCATGCGCCCTCGCAAGGATGACACGAAATTCGTGAAGCGTCCATATTA 58
 QY 378 AATAATAATAATAATAAGAGAAAGGAAAAATTAAGAAAAATA 426
 DB 57 AAGAAAAAAGAAAAAGAAACAAAAACCCCAACAAAAACAAACAAGA 9

RESULT 13
 BF738217
 LOCUS CM3-KT0033-151200-572-f07 KT0033 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF738217
 ACCESSION BF738217.1 GI:12064893
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 369)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=CM3&2=CM3-KT0033-151200-572-f07&t3=2000-12-15&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 15
 High quality sequence stop: 369.

FEATURES

source
 1. 369
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="KT0033"
 /note="Organ: bladder tumor; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196.716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT 90 a 85 c 78 g 116 t
 ORIGIN

Query Match 4.5%; Score 93.6; DB 10; Length 369;
 Best Local Similarity 67.3%; Pred. No. 1.8e+03;
 Matches 132; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
 QY 1135 TATATATAGTATTGTTGGTAGAGATGGGATTTTGGCATTTTCCCGAGGCTAGTATTGAAC 1194
 DB 79 TAATTTTGTATTGTTGGTAAGATGGGTTTCCACATTTTGGCAGGCTGCTTTGAAT 138
 QY 1195 TCTAAGCTAAGCAATCTTCCTGCTCTGCGCTCCCAAAATGTTGGATTACAGGTGTAAG 1254
 DB 139 TCTTGACCAAGTGATCTCCCTGCTCTGCGCTCCCAAAAGTGTGGATTACAGGTGTGA 198
 QY 1255 CCACGTGACCCGGCTGATAGTGGTTTCATTTTACTTCTTGTACCACTCTGATCCAT 1314
 DB 199 TCATACACCCAGCCTTTAGATTGTTGAAGTACATTTCTAGGTTTACCAATGATCA 258
 QY 1315 TTTGAAGTAAAAATGC 1330
 DB 259 AATTGCTTAAGATGC 274

RESULT 14
 BF829139/c
 LOCUS MRI-HN0070-151200-003-h04 HN0070 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF829139
 ACCESSION BF829139.1 GI:12174344
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 464)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL

MEDLINE PUBMED COMMENT	20202663 10737800 Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR1&t2=MR1-HN0070- 151200-003-h04&t3=2000-12-15&t4=1) Seq primer: puc 18 forward High quality sequence start: 17 High quality sequence stop: 464.	JOURNAL COMMENT	plasmid inserts Unpublished Contact: Robert B. Weiss University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0065 row: G column: 12 Seq primer: CACACAGGAACAGCTATGACC Class: plasmid ends High quality sequence stop: 723.
FEATURES source	Location/Qualifiers 1..464 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /dev_stage="Adult" /clone_lib="HN0070" /note="Organ: head normal; Vector: puc18; Site 1: Sma1; Site 2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions." 139 a 95 c 145 g 86 t	FEATURES Location/Qualifiers 1..723 /organism="Mus musculus" /mol_type="genomic DNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UUGC1M0065G12" /sex="Male" /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-" /clone_lib="Mouse 10kb plasmid UUGC1M library" /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi 4732114 gb AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance." 275 a 105 c 175 g 167 t 1 others	BASE COUNT 275 a 105 c 175 g 167 t 1 others ORIGIN Query Match 4.5%; Score 93.6; DB 28; Length 723; Best Local Similarity 71.1%; Pred. No. 1.2e+03; Matches 123; Conservative 0; Mismatches 50; Indels 0; Gaps 0; QY 240 AGATGAAGAGACCAATGAAATAGTAATGACTCTGTTTCTTCAGCAGGACATATACTAA 299 Db 530 AGACAAAGCACATAATTANAATGAGAGTACTGTGCTCGCTTCGGCAGCACATATACTAA 589 QY 300 AATAGGAGCTATACAAAGAAGATTAGCATGGACTCTGTGCAAGAATGACACACAAATTTG 359 Db 590 AATTGGAACGATACAGAGAAGATTAGCATGGCCCTCGCAGGATGACACGCAATTCG 649 QY 360 TGAACATTTCCATATATTAAATAATAATAATAAAGAGAAAGAGAAAAA 412 Db 650 TGAAGCGTTCCATATATTTTGAATACAAAAATGCAAAAGGATCTCTAACTCAAAA 702
RESULT 15 AZ335882 LOCUS	723 bp DNA linear GSS 29-SEP-2000 1M0065G12R Mouse 10kb plasmid UUGC1M library Mus musculus genomic Clone UUGC1M0065G12 R, genomic survey sequence.		
DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM	AZ335882 1M0065G12R Mouse 10kb plasmid UUGC1M library Mus musculus genomic Clone UUGC1M0065G12 R, genomic survey sequence. GSS. GI:10404635 Mus musculus (house mouse)		
REFERENCE AUTHORS TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 723) Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R. Mouse whole genome scaffolding with paired end reads from 10kb		Search completed: November 14, 2003, 06:44:49 Job time : 4319.25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 03:50:07 ; Search time 117.926 Seconds
(without alignments)
7762.738 Million cell updates/sec

Title: US-10-005-337A-2
Perfect score: 2074
Sequence: 1 ctgcagcagttacttaagt.....acaagactcttcagccaac 2074

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A-COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B-COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A-COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B-COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS-COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92.2	4.4	66804	US-09-740-041-3	Sequence 3, Appli
2	89.4	4.3	6669	US-09-212-971-5	Sequence 5, Appli
3	89.4	4.3	6669	US-08-800-929A-5	Sequence 5, Appli
4	89.4	4.3	6669	US-09-617-053A-5	Sequence 5, Appli
5	89.4	4.3	75395	US-09-984-890-3	Sequence 3, Appli
6	87.8	4.2	59065	US-09-813-817-3	Sequence 3, Appli
7	87.8	4.2	59065	US-09-978-197-3	Sequence 3, Appli
8	86.8	4.2	956	US-09-641-638-36	Sequence 36, Appli
9	86.8	4.2	20674	US-09-641-638-651	Sequence 651, App
10	86.2	4.2	99500	US-09-798-096-10	Sequence 10, Appl
11	86	4.1	29629	US-09-729-995-3	Sequence 3, Appli
12	85.8	4.1	72604	US-09-268-992-7	Sequence 7, Appli
13	85.8	4.1	72604	US-09-657-474-7	Sequence 7, Appli
14	85.8	4.1	75395	US-09-984-890-3	Sequence 3, Appli
15	85.2	4.1	90541	US-09-759-359A-3	Sequence 3, Appli
16	85	4.1	330	US-09-078-294-24	Sequence 24, Appl
17	84.6	4.1	4550	US-09-103-663-35	Sequence 35, Appl
18	84.4	4.1	72604	US-09-268-992-7	Sequence 7, Appli
19	84.4	4.1	72604	US-09-657-474-7	Sequence 7, Appli
20	84.2	4.1	1001	US-09-641-638-199	Sequence 199, App
21	84.2	4.1	1001	US-09-641-638-200	Sequence 200, App
22	84.2	4.1	3001	US-09-539-333D-142	Sequence 142, App
23	84.2	4.1	35060	US-08-814-095-7	Sequence 7, Appli
24	84.2	4.1	46718	US-09-816-093-3	Sequence 3, Appli
25	84	4.1	53332	US-09-801-861-3	Sequence 3, Appli
26	84	4.1	246240	US-08-724-394A-20	Sequence 20, Appl
27	84	4.1	246240	US-08-724-394A-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-09-740-041-3
; Sequence 3, Application US/09740041
; Patent No. 6562593

; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Karl et al

; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF

; FILE REFERENCE: CL001001

; CURRENT APPLICATION NUMBER: US/09/740,041

; CURRENT FILING DATE: 2000-12-20

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 66804

; TYPE: DNA

; ORGANISM: Human

US-09-740-041-3

Query Match 4.4%; Score 92.2; DB 4; Length 66804;
Best Local Similarity 73.3%; Pred. No. 3.4e-13;

Matches 118; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1135 TATATATAGTATTGTTGGTAGAGATGGGATTTTGGCCATGTTGCCAGGCTAGTATTGGAAC 1194

Db 40127 TAATTTTGTATTGTTGGTAGAGATGGGATTTTGGCCATGTTGCCAGGCTAGTATTGGAAC 40186

QY 1195 TCCTAAGCTAAGCAATCTCTGCTCTGCTCCCAAAATGTTGGGATTACAGGTGTAAG 1254

Db 40187 TCCTGACTTAGGATGATCCGCCACCTCGGCTCCCAAACTGTTGGGATTACAGCATAG 40246

QY 1255 CCACCTGACCCGCTCATAGCTGTTTCAATTTACTCTATT 1295

Db 40247 CCACCTGCTGCGCCCTTTTTCATCTGCTCTGCT 40287

RESULT 2

US-09-212-971-5

; Sequence 5, Application US/09212971B

; Patent No. 6107041

; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G

; APPLICANT: MacKenzie, Alexander E

; APPLICANT: Liston, Peter

; APPLICANT: Baird, Stephen

; APPLICANT: Tsang, Benjamin K

; APPLICANT: Pratt, Christine

; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPS AND

; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE

Db 1230 TAATTTTGTATTTTAGTAGAGATGGGTTTCACCATTTTGGCCAGGCTGGTTTGAAC 1289
QY 1195 TCCTAAGCT-AAGCAATCTTCCTGCTCTGCCCTCCCAAAATGTTGGGATTACAGGTGTA 1253
Db 1290 TCCTGACCTCAAGTGATCCACTTGTCTTGCCCTCCCAAAATGCTGGGATTACAGGCGTGA 1349
QY 1254 GCCACTGCACCGGCTGATAGCTGGTTTCAATTTACTCTATTTCTTGACCACTCTG 1308
Db 1350 GCCACTGCACGAGGAGCCCTCTGTTTATATCTCTTTTGGCTCTACAGT 1404

RESULT 5

US-09-984-890-3
; Sequence 3, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 75395
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(75395)
; OTHER INFORMATION: n = A,T,C or G
US-09-984-890-3

Query Match 4.3%; Score 89.4; DB 4; Length 75395;

Best Local Similarity 72.1%; Pred. No. 1.8e-12;
Matches 129; Conservative 1; Mismatches 48; Indels 1; Gaps 1;
QY 1090 CAGAGTCCTGCCATTGCTAATATCTGCTCTCATTTTBTCTATATATATATAGTATTG 1149
Db 13177 CAGGCACCTGCCACCATCTGCTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTGTATTT 13236
QY 1150 TGTAGAGATGGGATTTTGGCATGTTGCCATGTTGCCAGGCTAGTATTGAACTCCTAAGCTAAGCAA 1209
Db 13237 TAGTAGAGATGGGTTTACCATTGTTGACCAGGCTGCTTGAACCTCTGACCTCA-TGA 13295
QY 1210 TCTTCTGCTCTGCTCTCCAAATGTTGGGATTACAGGTGTAAGCCACTGCACCCGGC 1268
Db 13296 TCTGCCACCTCGGCCCTCCAAAGTCTGGGATTACAGGCGTGAGCCACTGCACCCGCC 13354

RESULT 6

US-09-813-817-3
; Sequence 3, Application US/09813817
; Patent No. 6340583
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001178
; CURRENT APPLICATION NUMBER: US/09/813,817
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-813-817-3

Query Match 4.2%; Score 87.8; DB 4; Length 59065;

Best Local Similarity 78.1%; Pred. No. 4.2e-12;
Matches 118; Conservative 0; Mismatches 32; Indels 1; Gaps 1;
QY 1135 TATATATAGTATTGTTGGTAGAGATGGGATTTTGCATGTTGCCAGGCTAGTATTGAAC 1194
Db 11804 TAATTTTATATTTTAGTAGAGATGGGTTTGCATGTTGCCAAGCTAGTCTCGAAC 11863
QY 1195 TCCTAAGCT-AAGCAATCTTCCTGCTCTGCCCTCCCAAAATGTTGGGATTACAGGTGTA 1253
Db 11864 TCCTGACCTCAAGTGATCCACCTCGCTCGCCCTCCCAAGTGTGGGATTACAGGTGTA 11923
QY 1254 GCCACTGCACCGGCTGATAGCTGGTTTCAT 1284
Db 11924 GCCACTGCACCTGGCTGGAGGAGTATCTT 11954

RESULT 7

US-09-978-197-3
; Sequence 3, Application US/09978197
; Patent No. 6403353
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001178DIV
; CURRENT APPLICATION NUMBER: US/09/978,197
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 09/813,817
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-978-197-3

Query Match 4.2%; Score 87.8; DB 4; Length 59065;

Best Local Similarity 78.1%; Pred. No. 4.2e-12;
Matches 118; Conservative 0; Mismatches 32; Indels 1; Gaps 1;
QY 1135 TATATATAGTATTGTTGGTAGAGATGGGATTTTGCATGTTGCCAGGCTAGTATTGAAC 1194
Db 11804 TAATTTTATATTTTAGTAGAGATGGGTTTGCATGTTGCCAAGCTAGTCTCGAAC 11863
QY 1195 TCCTAAGCT-AAGCAATCTTCCTGCTCTGCCCTCCCAAAATGTTGGGATTACAGGTGTA 1253
Db 11864 TCCTGACCTCAAGTGATCCACCTCGCTCGCCCTCCCAAGTGTGGGATTACAGGTGTA 11923
QY 1254 GCCACTGCACCGGCTGATAGCTGGTTTCAT 1284
Db 11924 GCCACTGCACCTGGCTGGAGGAGTATCTT 11954

RESULT 8

US-09-641-638-36
; Sequence 36, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07

```

; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 36
; LENGTH: 956
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 478
; OTHER INFORMATION: 10-349-47 : polymorphic base C or T
; NAME/KEY: misc.binding
; LOCATION: 458..477
; OTHER INFORMATION: 10-349-47.mis1, potential
; NAME/KEY: misc.binding
; LOCATION: 479..498
; OTHER INFORMATION: 10-349-47.mis2, potential complement
; NAME/KEY: primer.bind
; LOCATION: 432..451
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer.bind
; LOCATION: 829..848
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc.binding
; LOCATION: 466..490
; OTHER INFORMATION: 10-349-47 potential probe
; US-09-641-638-36

Query Match
Best Local Similarity 4.2%; Score 86.8; DB 4; Length 956;
Matches 118; Conservative 1; Mismatches 34; Indels 1; Gaps 1;

QY 1116 GPTCATTTTBTCAATATATATAGTATTTGTGGTAGAGATGGGATTTTGGCCATGTT 1175
Db 9 GCCACCATGTGGCCCTGGCTAATTTTGTATTTTGTAGAGATGGGTTTCGCCATATT 68

QY 1176 GCCCAGGCTAGTATTGAACCTCTAAGCT-AAGCATCTTCCTGCTCTCGCCTCCCAAAAT 1234
Db 69 GCCCAGGCTGTCCTTGAACCTCTGGGCTCAAGTAATCTGCCTCGCCTCCCAAAAGT 128

QY 1235 GTTGGGATTACAGGTGAAGCCACATGCACCCGGC 1268
Db 129 GCTGGGATTACGCTGTGAGCCACTGCACCTGGC 162

RESULT 9
US-09-641-638-651
; Sequence 651, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bouqueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 651
; LENGTH: 20674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1123..3123
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 3124..3297
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 3871..4072
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 5552..5633
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 5758..5880
; OTHER INFORMATION: exon 4
; NAME/KEY: exon
; LOCATION: 5996..6099
; OTHER INFORMATION: exon 5
; NAME/KEY: exon
; LOCATION: 6349..6509
; OTHER INFORMATION: exon 6
; NAME/KEY: exon
; LOCATION: 7379..7522
; OTHER INFORMATION: exon 7
; NAME/KEY: exon
; LOCATION: 8645..8854
; OTHER INFORMATION: exon 8
; NAME/KEY: exon
; LOCATION: 12254..12340
; OTHER INFORMATION: exon 9
; NAME/KEY: exon
; LOCATION: 12854..13023
; OTHER INFORMATION: exon 10
; NAME/KEY: exon
; LOCATION: 13308..13429
; OTHER INFORMATION: exon 11
; NAME/KEY: exon
; LOCATION: 16567..16667
; OTHER INFORMATION: exon 12
; NAME/KEY: exon
; LOCATION: 16775..16945
; OTHER INFORMATION: exon 13
; NAME/KEY: exon
; LOCATION: 17063..17554
; OTHER INFORMATION: exon 14
; NAME/KEY: misc.feature
; LOCATION: 17555..20674
; OTHER INFORMATION: 3'regulatory region
; NAME/KEY: allele
; LOCATION: 1128
; OTHER INFORMATION: 10-508-191 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1182
; OTHER INFORMATION: 10-508-245 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1559
; OTHER INFORMATION: 10-509-284 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1570
; OTHER INFORMATION: 10-509-295 : deletion of C
; NAME/KEY: allele
; LOCATION: 1827
; OTHER INFORMATION: 10-510-173 : variable motif ATTGA or TTTT
; NAME/KEY: allele
; LOCATION: 2048
; OTHER INFORMATION: 10-511-62 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 2323
; OTHER INFORMATION: 10-511-337 : insertion of T
```


Db 32693 ATTTTGTGTTTTATTAGAGTGGGTTTTCCACATGTTGCCAGGCTGGTCTCGAACT 32752
QY 1196 CTAAGCT-AAGCAATCTCTGCTCTGCCCTCCAAAATGTTGGATTACAGGTGTAAG 1254
Db 32753 CTTGACCTCAAGCATTCACCTGCTCGCTCCCAAGTGTGGATTACAGGCATGAG 32812
QY 1255 CCACCTGCACCCGGCTGATAGCTGGTTTCATTTACTCTAT 1293
Db 32813 CCACCTGCACCCGGCTCTCCCTTTTCATTTATGCTGT 32851

RESULT 11
US-09-729-995-3/c
; Sequence 3, Application US/09729995
; Patent No. 6426206
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000904
; CURRENT APPLICATION NUMBER: US/09/729,995
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 29629
; TYPE: DNA
; ORGANISM: Human
US-09-729-995-3

Query Match 4.1%; Score 86; DB 4; Length 29629;
Best Local Similarity 77.6%; Pred. No. 9.2e-12;
Matches 104; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1135 TATATATAGTATTCTGGTAGAGATGGGATTTTGGCATGTTGCCAGGCTAGTATTGAAC 1194
Db 16878 TAATTTTGTATTTTAGTAGAGATGGGTTTACACATGTTGGCAGGCTGCTCTTAAC 16819
QY 1195 TCCTAAGCTAAGCAATCTCTGCTCTGCCCTCCAAAATGTTGGATTACAGGTGTAAG 1254
Db 16818 TCCTAAGCTAAGCAATCTCTGCTCTGCCCTCCAAAATGTTGGATTACAGGTGTAAG 1254
QY 1255 CCACCTGCACCCGGC 1268
Db 16758 CCACCGCACCCGGC 16745

RESULT 12
US-09-268-992-7/c
; Sequence 7, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; EARLIER FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-268-992-7

Query Match 4.1%; Score 85.8; DB 4; Length 72604;
Best Local Similarity 72.5%; Pred. No. 1.5e-11;
Matches 111; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 1135 TATATATAGTATTCTGGTAGAGATGGGATTTTGGCATGTTGCCAGGCTAGTATTGAAC 1194
Db 57580 TTTTGTGATTTTAAATAGAGATGGGTTTCCACATGTTGGTCTGAGCTGCTCGAAC 57521
QY 1195 TCCTAAGCTAAGCAATCTCTGCTCTGCCCTCCAAAATGTTGGATTACAGGTGTAAG 1254
Db 57520 TCCTAAGCTAAGCAATCTCTGCTCTGCCCTCCAAAATGTTGGATTACAGGTGTAAG 1254
QY 1255 CCACCTGCACCCGGCTGATAGCTGGTTTCATTTA 1287
Db 57460 CCACACACCCGGCTGTATGAAGCATGTTTTTA 57428

RESULT 13
US-09-657-474-7/c
; Sequence 7, Application US/09657474
; Patent No. 6399762
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/657,474
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/268,992
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/106,056
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/088,312
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/078,044
; PRIOR FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7

Query Match 4.1%; Score 85.8; DB 4; Length 72604;
Best Local Similarity 72.5%; Pred. No. 1.5e-11;
Matches 111; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 1135 TATATATAGTATTCTGGTAGAGATGGGATTTTGGCATGTTGCCAGGCTAGTATTGAAC 1194
Db 57580 TTTTGTGATTTTAAATAGAGATGGGTTTCCACATGTTGGTCTGAGCTGCTCGAAC 57521
QY 1195 TCCTAAGCTAAGCAATCTCTGCTCTGCCCTCCAAAATGTTGGATTACAGGTGTAAG 1254
Db 57520 TCCTAAGCTAAGCAATCTCTGCTCTGCCCTCCAAAATGTTGGATTACAGGTGTAAG 1254
QY 1255 CCACCTGCACCCGGCTGATAGCTGGTTTCATTTA 1287

Db 57460 CCACCACCCGGCTGTATGAAGCATGTTTTTA 57428

RESULT 14

US-09-984-890-3/c
; Sequence 3, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 75395
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(75395)
; OTHER INFORMATION: n = A,T,C or G
US-09-984-890-3

Query Match 4.1%; Score 85.8; DB 4; Length 75395;
Best Local Similarity 78.6%; Pred. No. 1.5e-11;
Matches 114; Conservative 1; Mismatches 29; Indels 1; Gaps 1;
QY 1125 TBTCTATATATATATAGTATTGTGGTAGAGATGGGATTTTGCCATGTGCCAGGCT 1184
DB 44960 TAGTTTTTATTATTCTATTTTTTTTGTAGAGATGGCATCTTGCCATCTTGCCAGGCT 44901
QY 1185 AGTATTGAACCTCTAGCT-AAGCAATCTCTCTCTGCTCCCAAAATGTTGGGATT 1243
DB 44900 GGCTCAAATCTCTAGGCTCAAGCAATCTCTGCTCAGCTCCCAAAAGTGCTGGGATT 44841
QY 1244 ACAGGTGTAAGCACTGCACCCGGC 1268
DB 44840 ACAGGCGTGAGCCACACCCAGC 44816

RESULT 15

US-09-759-359A-3/c
; Sequence 3, Application US/09759359A
; Patent No. 6492153
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDEH, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001043
; CURRENT APPLICATION NUMBER: US/09/759,359A
; CURRENT FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 90541
; TYPE: DNA
; ORGANISM: Human
US-09-759-359A-3

Query Match 4.1%; Score 85.2; DB 4; Length 90541;
Best Local Similarity 76.0%; Pred. No. 2.2e-11;
Matches 117; Conservative 1; Mismatches 35; Indels 1; Gaps 1;
QY 1116 GTCTCATATATATATATAGTATTGTGTAGAGATGGGATTTTGCCATGTT 1175
DB 23174 GGCTAATTTTGGTTTTTATTATTATTATTAAAGTAGAGATGGGATTTGCGCATGTT 23115
QY 1176 GCCCAGGCTAGTATTGAACCTCTTAAGCT-AAGCAATCTTCTCTCTCTGCCCTCCCAAAAT 1234

Db 23114 GGCACGGCTGGTCTCAAACTCTTAGCTCAAGCAATCCACCTGCTTGGCTCCCAAAGT 23055
QY 1235 GTTGGATTACAGGTGTAAGCCCACTGCACCCGGC 1268
DB 23054 GCTGGGATTACAGGCGTGAGCCACACAGTGTCTGGC 23021

Search completed: November 14, 2003, 06:49:03
Job time : 120.926 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 03:51:42 ; Search time 613.964 Seconds
(without alignments)
11040.154 Million cell updates/sec

Title: US-10-005-337A-2

Perfect score: 2074

Sequence: 1 CTGACGACGAGTCTTAAATG.....ACAAGACTCTTCAGCAAC 2074

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTU5_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2071.6	99.9	2074	14	US-10-005-337A-2
2	367.8	17.7	2358	14	US-10-005-337A-1
3	144.4	7.0	1889	10	US-09-974-298-182
4	144.4	7.0	1889	12	US-10-240-965-158
5	144.4	7.0	1889	12	US-10-252-157-421
6	141.4	6.8	1325	14	US-10-084-817-217
7	129.4	6.2	184	9	US-09-923-876-5218
8	95.4	4.6	6040	14	US-09-764-860-1029
9	95.4	4.6	6040	14	US-10-074-095-1029
10	94.8	4.6	300000	14	US-10-262-552-33
11	93.8	4.5	119040	11	US-09-911-077A-19
12	93.2	4.5	32220	10	US-09-764-877-3933
13	92.6	4.5	555	12	US-10-027-632-95562
14	92.6	4.5	555	12	US-10-027-632-95563
15	92.6	4.5	555	13	US-10-027-632-95562
16	92.6	4.5	555	13	US-10-027-632-95563

17	92.2	4.4	6804	9	US-09-740-041-3	Sequence 3, Appli
18	92.2	4.4	6804	12	US-10-389-967-3	Sequence 3, Appli
19	91	4.4	461	11	US-09-918-995-35181	Sequence 35181, A
20	91	4.4	763	12	US-10-027-632-27383	Sequence 27383, A
21	91	4.4	763	12	US-10-027-632-27384	Sequence 27384, A
22	91	4.4	763	13	US-10-027-632-27383	Sequence 27383, A
23	91	4.4	763	13	US-10-027-632-27384	Sequence 27384, A
24	91	4.4	56737	10	US-09-782-378A-17	Sequence 17, Appli
25	90.4	4.4	19820	10	US-09-764-877-2713	Sequence 2713, Ap
26	90.2	4.3	701	12	US-10-027-632-113259	Sequence 113259
27	90.2	4.3	701	12	US-10-027-632-113260	Sequence 113260
28	90.2	4.3	701	13	US-10-027-632-113259	Sequence 113259
29	90.2	4.3	701	13	US-10-027-632-113260	Sequence 113260
30	90.2	4.3	705	12	US-10-027-632-268052	Sequence 268052
31	90.2	4.3	705	13	US-10-027-632-268053	Sequence 268053
32	90.2	4.3	705	13	US-10-027-632-268052	Sequence 268052
33	90.2	4.3	705	13	US-10-027-632-268053	Sequence 268053
34	90	4.3	625	12	US-10-027-632-102687	Sequence 102687
35	90	4.3	625	12	US-10-027-632-102688	Sequence 102688
36	90	4.3	625	12	US-10-027-632-102689	Sequence 102689
37	90	4.3	625	13	US-10-027-632-102687	Sequence 102687
38	90	4.3	625	13	US-10-027-632-102688	Sequence 102688
39	90	4.3	625	13	US-10-027-632-102689	Sequence 102689
40	90	4.3	715	12	US-10-027-632-110575	Sequence 110575
41	90	4.3	715	12	US-10-027-632-110576	Sequence 110576
42	90	4.3	715	13	US-10-027-632-110575	Sequence 110575
43	90	4.3	715	13	US-10-027-632-110576	Sequence 110576
44	89.8	4.3	86592	12	US-10-211-160-1	Sequence 1, Appli
45	89.8	4.3	86592	12	US-10-051-681A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-10-005-337A-2
; Sequence 2, Application US/10005337A
; Publication No. US2003039984A1
; GENERAL INFORMATION:
; APPLICANT: BENOIT, Patrick
; APPLICANT: SCHWARTZ, Bertrand
; APPLICANT: BRANELLER, Didier
; APPLICANT: CHIEN, Kenneth R.
; TITLE OF INVENTION: SEQUENCES UPSTREAM OF THE CARP GENE, VECTORS CONTAINING
; TITLE OF INVENTION: THEM AND USES THEREOF
; FILE REFERENCE: 03806.0530-00000
; CURRENT APPLICATION NUMBER: US/10/005,337A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/251,582
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2074
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-005-337A-2

Query Match	99.9%	Score	2071.6	DB	14	Length	2074
Best Local Similarity	100.0%	Pred. No.	0				
Matches	2074	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	CTGACGACGAGTCTTAAATGTTTTTGGCTCTGACGATCTCTGTAAATGAGAGCATTA	60				
Db	1	CTGACGACGAGTCTTAAATGTTTTTGGCTCTGACGATCTCTGTAAATGAGAGCATTA	60				
QY	61	GTCTTCTCCAATTCGAGGCGATGACAGCTCTGGATTTTCATATCCAGACCCCTTAAA	120				
Db	61	GTCTTCTCCAATTCGAGGCGATGACAGCTCTGGATTTTCATATCCAGACCCCTTAAA	120				
QY	121	CATCCACAGTCTTCTCCCAACACTTCTCTCTTAACCTCTCAGTTGGGTGAG	180				
Db	121	CATCCACAGTCTTCTCCCAACACTTCTCTCTTAACCTCTCAGTTGGGTGAG	180				

181 GCCTGGAAACAAAAGGCATACGAAATGGTAGAAAAGGTGTCATGACTACTTCTGACTTA 240
181 GCCTGGAAACAAAAGGCATACGAAATGGTAGAAAAGGTGTCATGACTACTTCTGACTTA 240
241 GATGAAGAGACCAATGAAATAGTAATGACTCTGTTGCTTCAGCAGGACATATATAA 300
241 GATGAAGAGACCAATGAAATAGTAATGACTCTGTTGCTTCAGCAGGACATATATAA 300
301 ATAGAGAGTATACAAAGAGATAGCATGGAATCTGTGCAAGAAATGACACAAATTTGT 360
301 ATAGAGAGTATACAAAGAGATAGCATGGAATCTGTGCAAGAAATGACACAAATTTGT 360
361 GAAACATTTCCATATATATAAATAAATAAATAAATAAAGAGAAAGAAAATTAAGAAG 420
361 GAAACATTTCCATATATATAAATAAATAAATAAATAAAGAGAAAGAAAATTAAGAAG 420
421 AAAATAGTGATAGCTGTGTCCTCAATCTCAAGAAAAGCCAGGAGATTTCTTTATTACCC 480
421 AAAATAGTGATAGCTGTGTCCTCAATCTCAAGAAAAGCCAGGAGATTTCTTTATTACCC 480
481 CTTTAAAGATAGAAATATTAGGAGACCGGAACATATATGATACAGGAGTACTGGGAGGTCC 540
481 CTTTAAAGATAGAAATATTAGGAGACCGGAACATATATGATACAGGAGTACTGGGAGGTCC 540
541 CTCTTTGTCAATGTTTGTCTTTGGGTGGGAGTGCATGCTCTTCAAGATTTTCAGAAAC 600
541 CTCTTTGTCAATGTTTGTCTTTGGGTGGGAGTGCATGCTCTTCAAGATTTTCAGAAAC 600
601 ACCATCCACTGACTGAGCAATTCAGGGGCAAGAGAGATGCGAGCAATTTGTTGATT 660
601 ACCATCCACTGACTGAGCAATTCAGGGGCAAGAGAGATGCGAGCAATTTGTTGATT 660
661 GGGTGAGTTTGGGAGAAATAGACACAAAGGTCAAACTTCTTAATTAACACTT 720
661 GGGTGAGTTTGGGAGAAATAGACACAAAGGTCAAACTTCTTAATTAACACTT 720
721 CCTCCATTCACAAATTCCTCTCCCATTTCTCTCTCTCTTTACTSAKARAACCC 780
721 CCTCCATTCACAAATTCCTCTCCCATTTCTCTCTCTCTTTACTSAKARAACCC 780
781 AGTTTTTCTGAACTATATAAATACCCAGTAGTTTACATAATTTACACTCAAGA 840
781 AGTTTTTCTGAACTATATAAATACCCAGTAGTTTACATAATTTACACTCAAGA 840
841 TTGAAACACAGAAATAGAGACCTTTTCAACCCCTTCGGGAAGCAAGTGCATTCCTCC 900
841 TTGAAACACAGAAATAGAGACCTTTTCAACCCCTTCGGGAAGCAAGTGCATTCCTCC 900
901 AGCCAGGTGCTCAAACTTTGATGCATCAGAAATCATCTGGGTGCTTTKAAATTCAGATG 960
901 AGCCAGGTGCTCAAACTTTGATGCATCAGAAATCATCTGGGTGCTTTKAAATTCAGATG 960
961 ATTCTTACAGATTACCAATAATCACTCAGAAATTCCTCGAGTGGGCGCAGGATCTGTA 1020
961 ATTCTTACAGATTACCAATAATCACTCAGAAATTCCTCGAGTGGGCGCAGGATCTGTA 1020
1021 TTTCTGACAACTCCACAGGTGATTCCTTTTCCCAACAGCATTTTGAGAACTTCAGCTCAA 1080
1021 TTTCTGACAACTCCACAGGTGATTCCTTTTCCCAACAGCATTTTGAGAACTTCAGCTCAA 1080
1081 TGACCTAATCAGAGTCTGCCATTTGCTAATATCTGGTCTCATTTTBTCTATATATA 1140
1081 TGACCTAATCAGAGTCTGCCATTTGCTAATATCTGGTCTCATTTTBTCTATATATA 1140
1141 TAGTATTTGGTGTAGAGATGGGATTTTGGCATTTTCCAGGCTAGTATTGAACCTCTAA 1200
1141 TAGTATTTGGTGTAGAGATGGGATTTTGGCATTTTCCAGGCTAGTATTGAACCTCTAA 1200
1201 GCTAAGCAATCTTCTGTCTCTGCTCCCTCCCAAAATGTTGGGATTTACAGGTGTAAGCACTG 1260
1201 GCTAAGCAATCTTCTGTCTCTGCTCCCTCCCAAAATGTTGGGATTTACAGGTGTAAGCACTG 1260

1261 CACCGGCTGATAGTGGTCTTCACTTACTCTATTTCTTGACCACTCTGATCATTTTGAA 1320
1261 CACCGGCTGATAGTGGTCTTCACTTACTCTATTTCTTGACCACTCTGATCATTTTGAA 1320
1321 GTAAAAATGCTCCAAATTTATGCTGTTTTAGAACACGGTAAGCATGTCATGCTGTAATG 1380
1321 GTAAAAATGCTCCAAATTTATGCTGTTTTAGAACACGGTAAGCATGTCATGCTGTAATG 1380
1381 GCCAGTGACATCATAAAAAGAGTGCAATTAAGTAACTGCTTCAATGCTTTATAATGATG 1440
1381 GCCAGTGACATCATAAAAAGAGTGCAATTAAGTAACTGCTTCAATGCTTTATAATGATG 1440
1441 GTAAAGTGGCATGTATGGGGCTTATTTAGCCACAGCATCACTCCAAAGAAATTCCAAACA 1500
1441 GTAAAGTGGCATGTATGGGGCTTATTTAGCCACAGCATCACTCCAAAGAAATTCCAAACA 1500
1501 GATATAGCAAGTGTCTTTAGGGCCAGATCCCTTCCCTCAGGCTGTTTACCCAGGAA 1560
1501 GATATAGCAAGTGTCTTTAGGGCCAGATCCCTTCCCTCAGGCTGTTTACCCAGGAA 1560
1561 TAGGATGCTGGGCAAGTTTCCCTAAGTGAAGTGTGATAAGTCTGCTTATCAGAAA 1620
1561 TAGGATGCTGGGCAAGTTTCCCTAAGTGAAGTGTGATAAGTCTGCTTATCAGAAA 1620
1621 GATATTAAGTGGGGTGTGATATGTAAGGCATCTACATTTCTTGATAGTGTATATGA 1680
1621 GATATTAAGTGGGGTGTGATATGTAAGGCATCTACATTTCTTGATAGTGTATATGA 1680
1681 AAGCTGACAAAGAAAAAGGCGAGTGTGTTGCTCAATGTCAACAGACAGCTGCTCCCT 1740
1681 AAGCTGACAAAGAAAAAGGCGAGTGTGTTGCTCAATGTCAACAGACAGCTGCTCCCT 1740
1741 GACTCTTGACAAATAGGATGACTTGCATTGCTGAGGATGTGATCACCAACCAAGGAATG 1800
1741 GACTCTTGACAAATAGGATGACTTGCATTGCTGAGGATGTGATCACCAACCAAGGAATG 1800
1801 GCCCTCTCACATTTCTCTGATTTACATATTCAGCAGGTTAGCTTGTCTCCCTCCC 1860
1801 GCCCTCTCACATTTCTCTGATTTACATATTCAGCAGGTTAGCTTGTCTCCCTCCC 1860
1861 TCTTCAGCTTCCAGACACTGAGTCTGGAATGAAATTCACCTGCTCTGAGTTGGCTCC 1920
1861 TCTTCAGCTTCCAGACACTGAGTCTGGAATGAAATTCACCTGCTCTGAGTTGGCTCC 1920
1921 TAATGGGGCGGAGTGTACTTCCGTTCCAGAGTTGGAAGATTATCTCACCGGCCCCA 1980
1921 TAATGGGGCGGAGTGTACTTCCGTTCCAGAGTTGGAAGATTATCTCACCGGCCCCA 1980
1981 GCTATATAAGCTGACCGGTGTGGAGGGCCAGCAGGGCCAACTCCAGGGATTCTCTTCCA 2040
1981 GCTATATAAGCTGACCGGTGTGGAGGGCCAGCAGGGCCAACTCCAGGGATTCTCTTCCA 2040
2041 CGACAGAAAAACATACAGACTCTTTCAGCCAAAC 2074
2041 CGACAGAAAAACATACAGACTCTTTCAGCCAAAC 2074

RESULT 2

US-10-005-337A-1

; Sequence 1, Application US/10005337A

; Publication No. US20030039984A1

; GENERAL INFORMATION:

; APPLICANT: BENOIT, Patrick

; APPLICANT: SCHWARTZ, Bertrand

; APPLICANT: BRANELLEC, Didier

; APPLICANT: CHEN, Kenneth R.

; TITLE OF INVENTION: SEQUENCES UPSTREAM OF THE CARP GENE, VECTORS CONTAINING

; TITLE OF INVENTION: THEM AND USES THEREOF

; FILE REFERENCE: 03806.0530-00000

; CURRENT APPLICATION NUMBER: US/10/005,337A

; CURRENT FILING DATE: 2001-12-07

; PRIOR APPLICATION NUMBER: US 60/251,582

; PRIOR FILING DATE: 2000-12-07

```

; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2358
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-005-337A-1

Query Match      17.7%; Score 367.8; DB 14; Length 2358;
Best Local Similarity 76.4%; Pred. No. 1.2e-83;
Matches 597; Conservative 0; Mismatches 162; Indels 22; Gaps 11;

Qy 1289 TCTATTCTTCGACACACTCTGATCCATTTTGAAGTAAATAATCTCTCAATATTATGCTGTT 1348
Db 1540 TCTCTCTGTGCATCACTTCGCGCCGCTTTTGGGGT--AGATCCTCTGATTAGCCTTCAGAT 1597
Qy 1349 TTAGAACACGGTAAAGCATGTCTATGCTTA---ATGGCCAGTGACATCATATAAAGAAAGT 1405
Db 1598 TTAGAACACGGTGAAGCCTGTGGTGCACATAATTATGGCCAGTGACCATAGATGATCAAAAGT 1657
Qy 1406 GCATTACTGAATGCTTTTCAATGTCTTATAATGATGCTGAAGTGGCATGTCTATGGGGCCTA 1465
Db 1658 GCATTACTGAATGCTTTTCAATTTCTCTTAATGCTGTACGATGCGATGTCTACAGGGCCAT 1717
Qy 1466 TTTAGC--CCAGACATCACTCCAAAGAAATCCAAACAGATATAGACAAGTGCCTTTTAGGGC 1524
Db 1718 TTTAGCTGCAGACATCACTCCAGAGAAATTCAAACAGATAGAGACAAGTGGCACCACAGAC 1777
Qy 1525 CCAGATCCCTTCCCTCAGGCTGTTTACCCAGGGAATAGGATGTCC---TGGGACAAGTT 1581
Db 1778 CCAATCTCTTCCCTTCGGGCTGATATATCCCCAGAAATAGGATGTGCCCAAAGCAACACTTC 1837
Qy 1582 TCCCTTAAGTGAAGTGTGTAAAGTCTGCTTTATCAGAAAGATATTTACTGGGGGTGTGATA 1641
Db 1838 CCAGCCAACTGGAGTGTGTAAGTCCAGTTATCAGAAAGATATGGCTGTAAAGTGTGATG 1897
Qy 1642 TGTAGGGGATCTACATTTCTTGATA--GGTAGTCAATATGAAGCTGACAAGAA--AAAA 1698
Db 1898 CACAGTGC--TTGCATTTTCTTGATACGTTAGTCAATATGAGAGCTGACAAAGAAGGAAAA 1955
Qy 1699 AGGCAGTGTGTGGTGCMAATGTCAAAGACAGCTGTCCCCGTGAC--TCTTGACAAATAGG 1757
Db 1956 AGACAGCGATGTGGTGCMAATATTAACAGGCGAGCTGTCCCCCTGGCTTCCCGATACGTTGG 2015
Qy 1758 ATGACTTTGCATTTGCTGAGCGATGTGATCAACCAAGAGGAATGGCCCTCTCACATTTCTT 1817
Db 2016 ATGACTCGCATTTGCTGAGCGGTGTGGTCACTGCCAAGAGGAATGACCTCTCACATTTCTT 2075
Qy 1818 CTTGANTCAATATTCAGCAGGGTTAGCTGTCTCTCCCTCTCCCTCTTTCACTTTCCACAGAC 1877
Db 2076 CTTGANTCGCATACGCCGCGG----CCAGCTTGTCTATCTCCCTCTTTGGGCTTCCCGAC 2130
Qy 1878 ACTGAGTCTGGAATGAAATTTCACTGCTCTGAGTTGGCTCTCTAATGGGGGCGGGAGTG 1937
Db 2131 ACTAAGTCTGGAATGAAATTTCACTGCTCTGAAATTTGGCCACTGTGTGGGGCGAGGGGTG 2190
Qy 1938 TTACTTCGGTTTCCAGGTTGGAAGATTATCTCACCGGCGCCAGCTATATAAGCTGACCG 1997
Db 2191 TGACTTGGCTTCCAGGCTGGAAGATTATCTCACCCAGCCCTAGCTATATAA--CGGGCTG 2249
Qy 1998 GTGTGAGGGGGCCAGCAGGGCCCAACTCCAGGGATTCCTTTC--CACGACGAAAAAATAC 2056
Db 2250 GTGTGAGGGGGCTCCACAGGGCCAGTTTCCAGGGGTTCATCCACAGAGAGAGAAAAATAG 2309
Qy 2057 A 2057
Db 2310 A 2310

RESULT 3
US-09-974-298-182
; Sequence 182, Application US/09974298
; Patent No. US20020156263A1

```


	Qy	Db	Qy	Db
1988	AAGCTG-ACCGGTGTGAGGGGCCACGAGGCCAACTCAGGGGATTCCTTCCACGACAG	2046		
61	AAGTCNACCGTNTGAGGGGCCAGAGGGCCAAACCCAGGGGANTCCTTCCACGACAG	120		
2047	AAAAACATACAAGACTCTCTTCAGGCCAAC	2074		
121	ANAAACATNCAAGACTCTTCAGGCCAAC	148		

RESULT 8

```

US-09-764-860-1029/c
; Sequence 1029, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1029
; LENGTH: 6040
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-1029

```

	Query Match	4.6%;	Score 95.4;	DB 9;	Length 6040;
	Best Local Similarity	72.8%;	Pred. No. 2.1e-13;		
	Matches 123;	Conservative 0;	Mismatches 46;	Indels 0;	Gaps 0;
Qy	1136	ATATATAGTATTTGCTGTACAGATGGGATTTTCCCATGTTGCCAGGCTAGTATTGAACT	1195		
Db	3417	ATTTTATATTTTTAGTGAGATGGGGTTTCGCCATATTTGGCCAGGCTGGTCTTGAAC	3358		
Qy	1196	CCTAAGCTAAGCAATCTTCTGTCTCTGCGCTCCCAAATGTTGGGATTACAGGTGTAAAGC	1255		
Db	3357	CCTGACCTCAGCGATCTGCTTGCTCAGCCTCTCATGTGCTGGGATTACAGGCGGTGAGC	3298		
Qy	1256	CATGCACCCGGTGATAGCTGGTTTCATTTACTCTATTTCTTGACCAAC	1304		
Db	3297	CATGCACCTGGCGGTTTTTTTTCTTTTCTTTTTTTTTTTTTTTTTTTTGGAGAC	3249		

RESULT 9

```

US-10-074-095-1029/c
; Sequence 1029, Application US/10074095
; Publication No. US20030077704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: 09/764,860
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29

```

```
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08

Query Match 4.6%; Score 95.4; DB 14; Length 6040;
Best Local Similarity 72.8%; Pred. No. 2.1e-13;
Matches 123; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1136 ATATATAGTATTGTGGTAGAGATGGGATTTGCCATGTTGCCAGGCTAGTATTGAACT 1195
Db 3417 ATTTTATATTTTATTTAGTGGAGATGGGTTTCCCATATTTGCCAGGCTGGTCTTGAAC 3358

QY 1196 CCTAAGCTAAGCAATCTTCTGCTGCTCCCAAAATGTTGGGATTACAGGTGTAAGC 1255
Db 3357 CCGTACCTCAGCGATCTGCTGCTCAGCTTCAATGCTGGGATTACAGCGTGAGC 3298

QY 1256 CACTGCACCCGGCTGATAGCTGGTTTCATTACTCTATTCTTTGACCAC 1304
Db 3297 CACTGCACCTGGCCGTTTTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3249

RESULT 10
US-10-262-552-33
; Sequence 33, Application US/10262552
; Publication No. US20030125289A1
; GENERAL INFORMATION:
; APPLICANT: Gelb, Bruce D.
; APPLICANT: Tartaglia, Marco
; TITLE OF INVENTION: NOONAN SYNDROME GENE
; FILE REFERENCE: 2420/1J859-US1
; CURRENT APPLICATION NUMBER: US/10/262,552
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 60/326,532
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 300000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(300000)
; OTHER INFORMATION: where n may be a or g or c or t/u, unknown or other
US-10-262-552-33

Query Match 4.6%; Score 94.8; DB 14; Length 300000;
Best Local Similarity 77.1%; Pred. No. 3.3e-12;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 04:10:57 ; Search time 535.347 Seconds
(without alignments)
10457.963 Million cell updates/sec

Title: US-10-005-337A-2

Perfect score: 2074

Sequence: 1 ctcgacgaagttacttaag.....acaagactcttcagccaac 2074

Scoring table:

OLIGO_NUC
Capex 60.0 , Gapex 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2467832

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database : N Geneseq_19Jun03.*

1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

C 1	22	1.1	41	AAI71558	Human mitotic cycl
C 2	21	1.0	33	AAI71555	Human mitotic cycl
C 3	20	1.0	41	AAI71557	Human mitotic cycl
C 4	19	0.9	33	ABA97055	Human 2-hydroxy ac
C 5	18	0.9	34	AAQ57179	Enzymatic RNA mole
C 6	18	0.9	34	AAZ56029	AU-rich element (A
C 7	18	0.9	34	AAH26599	Human tumour neuro
C 8	18	0.9	34	AAH46912	Tumour necrosis fa

C 9	18	0.9	34	24	AAH46934	Tumour necrosis fa
C 10	17	0.8	17	20	AAA22696	Integrin subunit b
C 11	17	0.8	17	20	AAA22697	Integrin subunit b
C 12	17	0.8	33	21	AAZ56027	Transcription temp
C 13	17	0.8	33	24	ABO77985	DNA topoisomerase
C 14	16	0.8	17	20	AAA22698	Integrin subunit b
C 15	16	0.8	17	20	AAA22708	Integrin subunit b
C 16	16	0.8	18	21	AAA22544	Antisense oligonuc
C 17	16	0.8	19	21	AAZ70595	Human biallelic ma
C 18	16	0.8	30	21	AAZ45408	PCR primer 3'RACE2
C 19	16	0.8	33	24	ABN87490	Human CAK1 antigen
C 20	16	0.8	41	24	ABV74646	Protein 9.24 probe
C 21	16	0.8	41	24	ABV74647	Human cell witheri
C 22	16	0.8	41	24	ABS56513	Human cell witheri
C 23	16	0.8	43	24	ABZ27757	Candida essential
C 24	16	0.8	43	24	ABZ27757	Human map-related
C 25	16	0.8	47	21	AAZ67996	Human map-related
C 26	16	0.8	47	21	AAZ68194	Human map-related
C 27	16	0.8	50	22	AAZ28423	Human SNP oligonuc
C 28	15	0.7	15	16	AAT56350	Mouse TNF-a hammer
C 29	15	0.7	15	16	AAT56332	Mouse TNF-a hammer
C 30	15	0.7	15	16	AAT56338	Mouse TNF-a hammer
C 31	15	0.7	15	16	AAT55815	Human TNF-alpha ha
C 32	15	0.7	15	16	AAT55817	Human TNF-alpha ha
C 33	15	0.7	15	16	AAT55819	Human TNF-alpha ha
C 34	15	0.7	15	16	AAT55799	Human TNF-alpha ha
C 35	15	0.7	15	16	AAT55801	Human TNF-alpha ha
C 36	15	0.7	15	22	AAF80980	PTGS2 allele speci
C 37	15	0.7	17	20	AAA22695	Integrin subunit b
C 38	15	0.7	17	20	AAA22707	Integrin subunit b
C 39	15	0.7	17	20	AAA22709	Integrin subunit b
C 40	15	0.7	17	20	AAA22898	Integrin subunit b
C 41	15	0.7	17	20	AAA22904	Integrin subunit b
C 42	15	0.7	17	25	ABZ61156	Human K-Ras DNazym
C 43	15	0.7	18	24	ABK43380	Siglec-BMS, PCR pr
C 44	15	0.7	18	24	ABK43392	Siglec-BMS, PCR pr
C 45	15	0.7	19	22	ABA82197	Zmax1 gene region

ALIGNMENTS

RESULT 1
AAI71558/c
ID AAI71558 standard; DNA; 41 BP.
XX
AC AAI71558;
XX
DT 10-JAN-2002 (first entry)
XX
DE Human mitotic cycle regulation protein 9 cDNA probe #2.
XX
KW Human; mitotic cycle regulation protein 9; cancer; haemopathy;
KW immunological disease; HIV infection; inflammation; gene therapy;
KW probe; ss.
XX
OS Homo sapiens.
XX
PN WO200175033-A2.
XX
PD 11-OCT-2001.
XX
PF 23-MAR-2001; 2001WO-CN00398.
XX
PR 24-MAR-2000; 2000CN-0115092.
XX
PA (SHAN-) SHANGHAI BIO WINDOW GENE DEV INC.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2001-626411/72.
XX
PT Human mitotic cycle regulation protein 9 and encoded polynucleotide,

PT used in diagnosis and treatment of malignant tumors, hemopathy, human
PT immunodeficiency virus infection, immunological diseases and
PT inflammation -
XX
XX
XX Example 6; Page 21; 33pp; Chinese.
XX The present invention provides the protein and coding sequences of human
CC mitotic cycle regulation protein 9. The sequences can be used in the
CC treatment of cancer, haemopathy, HIV infection, immunological diseases
CC and inflammation. The present sequence is a probe for the coding
CC sequence of the invention.
XX
XX Sequence 41 BP; 17 A; 12 C; 3 G; 9 T; 0 other;
SQ
Query Match 1.1%; Score 22; DB 22; Length 41;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1150 TGTAGAGATGGGATTTTGCCA 1171
DB 22 TGTAGAGATGGGATTTTGCCA 1

RESULT 2
AAI71555/c
ID AAI71555 standard; DNA; 33 BP.
XX
AC AAI71555;
XX
DT 10-JAN-2002 (first entry)
XX
DE Human mitotic cycle regulation protein 9 cDNA PCR primer #3.
XX
XX Human; mitotic cycle regulation protein 9; cancer; haemopathy;
KW immunological disease; HIV infection; inflammation; gene therapy;
KW PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200175033-A2.
XX
PD 11-OCT-2001.
XX
PF 23-MAR-2001; 2001WO-CN00398.
XX
PR 24-MAR-2000; 2000CN-0115092.
XX
PA (SHAN-) SHANGHAI BIO WINDOW GENE DEV INC.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2001-626411/72.
XX
XX Human mitotic cycle regulation protein 9 and encoded polynucleotide,
KW used in diagnosis and treatment of malignant tumors, hemopathy, human
KW PCR primer; ss.
XX
XX Homo sapiens.
XX
PN WO200175033-A2.
XX
PD 11-OCT-2001.
XX
PF 23-MAR-2001; 2001WO-CN00398.
XX
PR 24-MAR-2000; 2000CN-0115092.
XX
PA (SHAN-) SHANGHAI BIO WINDOW GENE DEV INC.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2001-626411/72.
XX
XX Human mitotic cycle regulation protein 9 and encoded polynucleotide,
PT used in diagnosis and treatment of malignant tumors, hemopathy, human
PT immunodeficiency virus infection, immunological diseases and
PT inflammation -
XX
XX
XX Example 4; Page 19; 33pp; Chinese.
XX The present invention provides the protein and coding sequences of human
CC mitotic cycle regulation protein 9. The sequences can be used in the
CC treatment of cancer, haemopathy, HIV infection, immunological diseases
CC and inflammation. The present sequence is a PCR primer for the coding
CC sequence of the invention.
XX
XX Sequence 33 BP; 12 A; 11 C; 3 G; 7 T; 0 other;
SQ
Query Match 1.0%; Score 21; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1152 GTAGAGATGGGATTTTGCCA 1172

DB 30 GTAGAGATGGGATTTTGCCA 10

RESULT 3
AAI71557/c
ID AAI71557 standard; DNA; 41 BP.
XX
AC AAI71557;
XX
DT 10-JAN-2002 (first entry)
XX
DE Human mitotic cycle regulation protein 9 cDNA probe #1.
XX
XX Human; mitotic cycle regulation protein 9; cancer; haemopathy;
KW immunological disease; HIV infection; inflammation; gene therapy;
KW probe; ss.
XX
OS Homo sapiens.
XX
PN WO200175033-A2.
XX
PD 11-OCT-2001.
XX
PF 23-MAR-2001; 2001WO-CN00398.
XX
PR 24-MAR-2000; 2000CN-0115092.
XX
PA (SHAN-) SHANGHAI BIO WINDOW GENE DEV INC.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2001-626411/72.
XX
XX Human mitotic cycle regulation protein 9 and encoded polynucleotide,
PT used in diagnosis and treatment of malignant tumors, hemopathy, human
PT immunodeficiency virus infection, immunological diseases and
PT inflammation -
XX
XX Example 6; Page 21; 33pp; Chinese.
XX The present invention provides the protein and coding sequences of human
CC mitotic cycle regulation protein 9. The sequences can be used in the
CC treatment of cancer, haemopathy, HIV infection, immunological diseases
CC and inflammation. The present sequence is a probe for the coding
CC sequence of the invention.
XX
XX Sequence 41 BP; 18 A; 11 C; 3 G; 9 T; 0 other;
SQ
Query Match 1.0%; Score 20; DB 22; Length 41;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1152 GTAGAGATGGGATTTTGCCA 1171
DB 20 GTAGAGATGGGATTTTGCCA 1

RESULT 4
ABA97055/c
ID ABA97055 standard; DNA; 33 BP.
XX
AC ABA97055;
XX
DT 18-MAR-2002 (first entry)
XX
DE Human 2-hydroxy acid dehydrogenase 16 PCR primer SEQ ID 6.
XX
XX Human; 2-hydroxy acid dehydrogenase 16; cancer; haemopathy; HIV;
KW human immunodeficiency virus; infection; PCR primer; ss.
XX
OS Homo sapiens.
XX

PN CN1315516-A.
 XX
 PD 03-OCT-2001.
 XX
 PF 24-MAR-2000; 2000CN-0115096.
 XX
 PR 24-MAR-2000; 2000CN-0115096.
 XX
 PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
 XX
 PI Mao Y, Xie Y;
 XX
 DR WPI; 2002-056337/08.
 XX
 PT New human 2-hydroxy acid dehydrogenase 16 and encoding polynucleotide,
 PT useful for treating cancer, haemopathy and human immunodeficiency virus
 PT infection -
 XX
 PS Example 4; Page 19 (Disclosure); 33pp; Chinese.
 XX
 CC This invention describes a novel human 2-hydroxy acid dehydrogenase 16,
 CC its recombinant production, its encoding polynucleotide and application
 CC and a 2-hydroxy acid dehydrogenase 16 antagonist. The polypeptide is
 CC useful for treating cancer, haemopathy and human immunodeficiency virus
 CC infection. This sequence represents a PCR primer used in the
 CC amplification of the human 2-hydroxy acid dehydrogenase 16 described in.
 CC the method of the invention.
 XX
 SQ Sequence 33 BP; 8 A; 11 C; 4 G; 10 T; 0 other;
 XX
 Query Match 0.9%; Score 19; DB 24; Length 33;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1230 AAAATGTTGGATTACAGG 1248
 DB 33 AAAATGTTGGATTACAGG 15
 RESULT 5
 AAQ57179/c
 ID AAQ57179 standard; mRNA; 34 BP.
 AC AAQ57179;
 XX
 DT 25-MAR-2003 (updated)
 DT 26-JUL-1994 (first entry)
 XX
 DE Enzymatic RNA molecule TNF-alpha mRNA target sequence.
 XX
 KW Tumour necrosis factor; specific; cleavage; target RNA; protein;
 KW expression; inhibitor; inhibition; ribozyme; treatment; prophylaxis;
 KW prevention; psoriasis; asthma; inflammatory diseases; restenosis;
 KW cardiovascular condition; hypertension; arthritis; ss.
 XX
 OS Synthetic.
 XX
 WO9402595-A1.
 XX
 PD 03-FEB-1994.
 XX
 PF 02-JUL-1993; 93WO-US06316.
 XX
 PR 17-JUL-1992; 92US-0916763.
 PR 07-DEC-1992; 92US-0987132.
 PR 07-DEC-1992; 92US-0989948.
 PR 07-DEC-1992; 92US-0989949.
 PR 19-JAN-1993; 93US-0008895.
 XX
 PA (RIBO-) RIBOZYME PHARM INC.
 XX
 PI Draper KG, Sullivan SM;
 XX

DR WPI; 1994-048853/06.
 XX
 PT Enzymatic RNA molecules which cleave mRNA - used to treat or
 PT prevent inflammatory, arthritic, stenotic or cardiovascular diseases
 PT or conditions
 XX
 PS Claim 3; Page 16; 65pp; English.
 XX
 CC This is a TNF-alpha mRNA target sequence (nucleotide no. 1344) of an
 CC enzymatic RNA molecule (ribozyme) which cleaves mRNA associated with
 CC the development or maintenance of a peoriatic or asthmatic condition.
 CC The concn. of the ribozyme necessary to effect a therapeutic treatment
 CC is lower than that of an antisense oligonucleotide and the specificity
 CC of action is higher.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 34 BP; 10 A; 0 C; 0 G; 24 T; 0 other;
 XX
 Query Match 0.9%; Score 18; DB 15; Length 34;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 380 AAATAAATAAATAATAAA 397
 DB 33 AAATAAATAAATAATAAA 16
 RESULT 6
 AAZ56029/c
 ID AAZ56029 standard; RNA; 34 BP.
 XX
 AC AAZ56029;
 XX
 DT 23-MAR-2000 (first entry)
 XX
 DE AU-rich element (ARE) sequence #1.
 XX
 KW AU rich element; ARE; adenylation; differentiation; rheumatoid arthritis;
 KW transcription template; inflammatory bowel disease; ss.
 XX
 OS Synthetic.
 XX
 WO9961605-A2.
 XX
 PD 02-DEC-1999.
 XX
 PF 26-MAY-1999; 99WO-US11581.
 XX
 PR 26-MAY-1998; 98US-0086675.
 XX
 PA (UYNE-) UNIV NEW JERSEY.
 XX
 PI Wilusz J, Ford LP;
 XX
 DR WPI; 2000-086719/07.
 XX
 PT In vitro system for studying regulated RNA turnover, containing cell
 PT extract and target RNA, for identifying modulators of RNA stability,
 PT potential therapeutic agents -
 XX
 PS Disclosure; Page 53; 80pp; English.
 XX
 CC This is an AU-rich element (ARE) sequence used in competition studies to
 CC evaluate the role of ARE proteins in the process of RNA deadenylation or
 CC degradation. It is thought that AREs influence mRNA degradation rates.
 CC The invention relates to an in vitro system (consisting of a cell extract
 CC and an exogenous target RNA) that is able to recapitulate regulated RNA
 CC turnover of the target RNA. The system is used to identify agents that
 CC modulate stability, deadenylation or degradation of the target RNA, or
 CC endogenous molecules that participate in deadenylation/degradation of the
 CC target RNA. These modulators e.g AREs are preferably involved in cell
 CC growth and differentiation in mammals, especially where these processes
 CC are implicated in cell transformation and immune system dysfunction and

CC are potential therapeutic agents, e.g. in conditions associated with
 CC abnormal expression of tumour necrosis factor- α . Examples of such
 CC diseases include sepsis, rheumatoid arthritis or inflammatory bowel
 CC disease. The system can also be used diagnostically to detect the
 CC molecular defects in such conditions and for development of improved gene
 CC delivery systems.

XX Sequence 34 BP; 10 A; 0 C; 0 G; 24 U; 0 other;

Query Match 0.9%; Score 18; DB 21; Length 34;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 AAATAAATAAATAATAAA 397
 |||||
 DB 33 AAATAAATAAATAATAAA 16

RESULT 7

AAH26599/c
 ID AAH26599 standard; mRNA; 34 BP.

XX AC AAH26599;

XX DT 12-NOV-2001 (first entry)

XX DE Human tumour necrosis factor gene 3' UTR AU-rich element.

XX KW Tumour necrosis factor; human; AU-rich element; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 FT misc_feature 4..8

FT /tag= a

FT /note= "AUUUA motif"

FT 8..11

FT /tag= b

FT /note= "AUUUA motif"

FT 11..15

FT /tag= c

FT /note= "AUUUA motif"

FT 15..19

FT /tag= d

FT /note= "AUUUA motif"

FT 22..26

FT /tag= e

FT /note= "AUUUA motif"

FT 26..30

FT /tag= f

FT /note= "AUUUA motif"

FT 30..34

FT /tag= g

FT /note= "AUUUA motif"

XX WO200164921-A1.

XX PD 07-SEP-2001.

XX PF 28-FEB-2001; 2001WO-US06782.

XX PR 29-FEB-2000; 2000US-0515369.

XX PA (UYCO) UNIV COLUMBIA NEW YORK.

XX PI Fisher PB, Madireddi MT;

XX WPI; 2001-565508/63.

XX Melanoma differentiation associated gene-7 promoter capable of

PT treating cancer comprises directing transcription of heterologous

PT coding sequence encoding tumour suppressor polypeptide positioned

PT downstream, useful for treating cancer

XX Disclosure; Fig 2C; 132pp; English.

XX The present sequence is that of an AU-rich sequence in the 3'
 CC untranslated region (3'UTR) of human tumour necrosis factor mRNA.
 CC The presence of AU-rich elements (AREs) in eukaryotic mRNAs
 CC correlates with rapid mRNA turnover and post-translational control.
 CC The ARE consists of multiple AUUUA motifs or sequences resembling
 CC it. A similar ARE sequence is found in the 3' UTR of the human
 CC melanoma differentiation associated gene-7 (Mda-7) gene (see
 CC AAH26596). The invention provides recombinant expression constructs
 CC in which the human Mda-7 promoter (see AAH26595) is operably linked
 CC to a coding sequence encoding a tumour suppressor protein. A
 CC pharmaceutical composition including the recombinant expression
 CC construct is used in a claimed method of treating melanoma,
 CC neuroblastoma, astrocytoma, glioblastoma multiforme, cervical
 CC cancer, breast cancer, colon cancer, prostate cancer, osteosarcoma,
 CC chondrosarcoma or a cancer of the central nervous system.

XX Sequence 34 BP; 10 A; 0 C; 0 G; 24 U; 0 other;

Query Match 0.9%; Score 18; DB 22; Length 34;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 AAATAAATAAATAATAAA 397

|||||
 DB 33 AAATAAATAAATAATAAA 16

RESULT 8

AAH26599/c

ID AAD46912 standard; DNA; 34 BP.

XX AC AAD46912;

XX DT 27-JAN-2003 (first entry)

XX DE Tumour necrosis factor (TNF) AU-rich element (ARE) DNA.

XX Gene expression; transcript stability; drug screening; AU-rich element;

XX KW ARE; tumour necrosis factor; TNF; ds.

XX OS Mammalia.

XX PN WO200272844-A1.

XX PD 19-SEP-2002.

XX PF 08-MAR-2002; 2002WO-AU00351.

XX PR 09-MAR-2001; 2001US-274770P.

XX PA (GENE-) GENE STREAM PTY LTD.

XX PI Daly J;

XX WPI; 2002-759847/82.

XX New expression vector useful for modulating gene expression,

PT identifying and analyzing regulatory sequences, new targets and

PT reagents for treating human diseases, comprises a transcribable

PT polynucleotide encoding an RNA element

XX Claim 11; Page 65; 103pp; English.

XX The present invention relates to novel expression vectors and/or reporter

CC vectors providing kinetics of protein expression with improved temporal

CC correlation to the promoter activity. The expression vectors comprise

CC transcribable polynucleotides having sequences of nucleotides encoding

CC RNA elements which modulates the stability of a transcript corresponding

CC to the transcribable polynucleotide. The expression vectors are useful

CC for modulating the stability of a transcript and determining expression

CC of a polynucleotide of interest. They are useful for modulating gene
 CC expression, identifying and analysing regulatory sequences, new targets
 CC and reagents for treating human diseases and for drug screening. The
 CC present sequence is tumour necrosis factor (TNF) ARE (AU-rich element)
 CC DNA. This sequence is used in the exemplification of the invention.
 XX
 SQ Sequence 34 BP; 10 A; 0 C; 0 G; 24 T; 0 other;

Query Match 0.9%; Score 18; DB 24; Length 34;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 AAATAAATAAATAAATAA 397
 DB 33 AAATAAATAAATAAATAA 16

RESULT 9
 AAD46934/C

ID AAD46934 standard; RNA; 34 BP.

XX AAD46934;

XX 27-JAN-2003 (first entry)

XX Tumour necrosis factor (TNF) AU-rich element (ARE) RNA.

XX Gene expression; transcript stability; drug screening; AU-rich element;
 KW ARE; tumour necrosis factor; TNF; ss.

XX Mammalia.

XX WO200272844-A1.

XX 19-SEP-2002.

XX 08-MAR-2002; 2002WO-AU00351.

XX 09-MAR-2001; 2001US-274770P.

XX (GENE-) GENE STREAM PTY LTD.

XX Daly J;

XX WPI; 2002-759847/82.

XX New expression vector useful for modulating gene expression,
 PT identifying and analysing regulatory sequences, new targets and
 PT reagents for treating human diseases, comprises a transcribable
 PT polynucleotide encoding an RNA element

XX Example 15; Page 65; 103pp; English.

XX The present invention relates to novel expression vectors and/or reporter
 CC vectors providing kinetics of protein expression with improved temporal
 CC correlation to the promoter activity. The expression vectors comprise
 CC transcribable polynucleotides having sequences of nucleotides encoding
 CC RNA elements which modulates the stability of a transcript corresponding
 CC to the transcribable polynucleotide. The expression vectors are useful
 CC for modulating the stability of a transcript and determining expression
 CC of a polynucleotide of interest. They are useful for modulating gene
 CC expression, identifying and analysing regulatory sequences, new targets
 CC and reagents for treating human diseases and for drug screening. The
 CC present sequence is tumour necrosis factor (TNF) ARE (AU-rich element)
 CC RNA. This sequence is used in the exemplification of the invention.

XX Sequence 34 BP; 10 A; 0 C; 0 G; 24 U; 0 other;

Query Match 0.9%; Score 18; DB 24; Length 34;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 AAATAAATAAATAAATAA 397

Db 33 AAATAAATAAATAAATAA 16

RESULT 10

AA22696/C

ID AA22696 standard; RNA; 17 BP.

XX AA22696;

XX 19-JUN-2000 (first entry)

XX Integrin subunit beta 3 substrate sequence SEQ ID NO:5922.

XX Human; aryl hydrocarbon nuclear transport; ARNT; Tie-2; angiogenesis;
 KW integrin alpha 6 subunit; integrin subunit beta 3; hairpin ribozyme;
 KW hammerhead ribozyme; angiogenic factor; cytosolic; antidiabetic;
 KW ophthalmologic; antiinflammatory; antiarthritic; antiporistic; ARMD;
 KW dermatological; RNA cleavage; cancer; diabetic retinopathy; arthritis;
 KW age related macular degeneration; inflammation; neovascular glaucoma;
 KW myopic degeneration; psoriasis; verruca vulgaris; angiofibroma;
 KW tuberosus sclerosis; pot-wine stain; Sturge Weber syndrome;
 KW Kippel-Trenaunay-Weber syndrome; Osler-Weber-Rendu syndrome; ss.

XX Homo sapiens.

XX WO9950403-A2.

XX 07-OCT-1999.

XX 24-MAR-1999; 99WO-US06507.

XX 27-MAR-1998; 98US-0079678.

XX (RIBO-) RIBOZYME PHARM INC.

XX Pavco PA, Roberts E, Jarvis T, Coeshott C, McSwiggen JA;

XX WPI; 1999-591315/50.

XX Novel ribozymes for modulating the synthesis, expression and/or
 PT stability of an mRNA encoding an angiogenic factors

XX Claim 54; Page 236; 305pp; English.

XX The present invention describes enzymatic nucleic acid molecules with
 CC RNA cleaving activity, which specifically cleave RNA encoded by an aryl
 CC hydrocarbon nuclear transporter (ARNT) gene, an integrin subunit beta 3
 CC gene, an integrin alpha 6 subunit gene, or a Tie-2 gene. AAAL6775 to
 CC AAAL1767 and AAAL17561 to AAAL17622 represent ribozyme sequences for ARNT,
 CC and AAAL1768 to AAAL17560 and AAAL17623 to AAAL17684 represent their
 CC corresponding target sequences; AAAL17685 to AAAL18385 and AAAL19087 to
 CC AAAL19154 represent ribozyme sequences for Tie-2, and AAAL18386 to AAAL19086
 CC AAAL19223 to AAAL19222 represent their corresponding target sequences;
 CC sequences for integrin alpha 6 subunit, and AAAL20362 to AAAL21500 and
 CC AAAL21596 to AAAL21688 represent their corresponding target sequences;
 CC AAAL21689 to AAAL22475 and AAAL23263 to AAAL23342 represent ribozyme
 CC for integrin subunit beta 3, and AAAL24476 to AAAL23262, AAAL23343 to
 CC AAAL34422 represent their corresponding target sequences. The ribozymes of
 CC the invention are used for modulating the synthesis, expression and/or
 CC stability of an mRNA encoding angiogenic factor, especially ARNT,
 CC integrin subunit beta-3, integrin subunit alpha-6, or Tie-2. They are
 CC especially used to treat cancer, diabetic retinopathy, age related
 CC macular degeneration (ARMD), inflammation, and arthritis, as well as
 CC neovascular glaucoma, myopic degeneration, psoriasis, verruca vulgaris,
 CC angiofibroma of tuberosus sclerosis, pot-wine stains, Sturge Weber
 CC syndrome, Kippel-Trenaunay-Weber syndrome, Osler-Weber-Rendu syndrome,
 CC and other syndromes and diseases related to the levels of ARNT, Tie-2,
 CC integrin subunit alpha-6, or integrin subunit beta-3.

XX Sequence 17 BP; 4 A; 0 C; 0 G; 13 U; 0 other;

SQ

Query Match 0.8%; Score 17; DB 20; Length 17;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 AATAAATAAATAATAA 397
 Db 17 AATAAATAAATAATAA 1

RESULT 11
 AAA22697/c
 ID AAA22697 standard; RNA; 17 BP.
 XX
 AC AAA22697;
 XX
 DT 19-JUN-2000 (first entry)
 XX
 DE Integrin subunit beta 3 substrate sequence SEQ ID NO:5923.
 XX
 KW Human; aryl hydrocarbon nuclear transport; ARNT; TIE-2; angiogenesis;
 KW integrin alpha 6 subunit; integrin subunit beta 3; hairpin ribozyme;
 KW hammerhead ribozyme; angiogenic factor; cytostatic; antidiabetic;
 KW ophthalmologic; antiinflammatory; antiarthritic; antipsoriatic; ARMD;
 KW dermatologic; RNA cleavage; cancer; diabetic retinopathy; arthritis;
 KW age related macular degeneration; inflammation; neovascular glaucoma;
 KW myopic degeneration; psoriasis; verruca vulgaris; angiofibroma;
 KW tuberos scleriosis; pot-wine stain; Sturge Weber syndrome;
 KW Kippel-Trenaunay-Weber syndrome; Osler-Weber-Rendu syndrome; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9950403-A2.
 XX
 PD 07-OCT-1999.
 XX
 PF 24-MAR-1999; 99WO-US06507.
 XX
 PR 27-MAR-1998; 98US-0079678.
 XX
 PA (RIBO-) RIBOZYME PHARM INC.
 XX
 PI Pavco PA, Roberts E, Jarvis T, Coeshott C, McSwiggen JA;
 XX
 DR WPI; 1999-591315/50.
 XX
 PT Novel ribozymes for modulating the synthesis, expression and/or
 PT stability of an mRNA encoding an angiogenic factors -
 XX
 PS Claim 54; Page 236; 305pp; English.
 XX

CC The present invention describes enzymatic nucleic acid molecules with
 CC RNA cleaving activity, which specifically cleave RNA encoded by an aryl
 CC hydrocarbon nuclear transporter (ARNT) gene, an integrin subunit beta 3
 CC gene, an integrin alpha 6 subunit gene, or a Tie-2 gene. AAAL6775 to
 CC AAAL1767 and AAAL17561 to AAAL17622 represent ribozyme sequences for ARNT,
 CC and AAAL1768 to AAAL17560 and AAAL17623 to AAAL17684 represent their
 CC corresponding target sequences; AAAL17685 to AAAL18385 and AAAL19087 to
 CC AAAL19154 represent ribozyme sequences for Tie-2, and AAAL18386 to AAAL19086
 CC and AAAL19155 to AAAL19222 represent their corresponding target sequences;
 CC AAAL19223 to AAAL20361 and AAAL21501 to AAAL21595 represent ribozyme
 CC sequences for integrin alpha 6 subunit, and AAAL20362 to AAAL21500 and
 CC AAAL1596 to AAAL21688 represent their corresponding target sequences;
 CC AAAL1689 to AAAL22475 and AAAL23263 to AAAL23342 represent ribozyme sequence
 CC for integrin subunit beta 3, and AAAL22476 to AAAL23262, AAAL23343 to
 CC AAAL23422 represent their corresponding target sequences. The ribozymes of
 CC the invention are used for modulating the synthesis, expression and/or
 CC stability of an mRNA encoding angiogenic factor, especially ARNT.
 CC integrin subunit beta-3, integrin subunit alpha-6, or Tie-2. They are
 CC especially used to treat cancer, diabetic retinopathy, age related
 CC macular degeneration (ARMD), inflammation, and arthritis, as well as
 CC neovascular glaucoma, myopic degeneration, psoriasis, verruca vulgaris,
 CC angiofibroma of tuberos scleriosis, pot-wine stains, Sturge Weber
 CC syndrome, Kippel-Trenaunay-Weber syndrome, Osler-Weber-Rendu syndrome,

CC and other syndromes and diseases related to the levels of ARNT, Tie-2,
 CC integrin subunit alpha-6, or integrin subunit beta-3.
 CC
 SQ Sequence 17 BP; 4 A; 0 C; 0 G; 13 U; 0 other;
 Query Match 0.8%; Score 17; DB 20; Length 17;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 AATAAATAAATAATAA 396
 Db 17 AATAAATAAATAATAA 1

RESULT 12
 AAZ56027/c
 ID AAZ56027 standard; DNA; 33 BP.
 XX
 AC AAZ56027;
 XX
 DT 23-MAR-2000 (first entry)
 XX
 DE Transcription template sequence for AU-rich element SVARE-A0.
 XX
 KW AU rich element; ARE; adenylation; differentiation; rheumatoid arthritis;
 KW transcription template; inflammatory bowel disease; ss.
 XX
 OS Synthetic.
 XX
 PN WO9961605-A2.
 XX
 PD 02-DEC-1999.
 XX
 PF 26-MAY-1999; 99WO-US11581.
 XX
 PR 26-MAY-1998; 98US-0086675.
 XX
 PA (UYNE-) UNIV NEW JERSEY.
 XX
 PI Wilusz J, Ford LP;
 XX
 DR WPI; 2000-086719/07.
 XX
 PT In vitro system for studying regulated RNA turnover, containing cell
 PT extract and target RNA, for identifying modulators of RNA stability,
 PT potential therapeutic agents -
 XX
 PS Disclosure; Page 53; 80pp; English.
 XX

CC This is a transcription template sequence for an AU rich element (ARE)
 CC protein. The synthetic oligonucleotide is used in evaluation of the role
 CC of ARE proteins in the process of RNA deadenylation or degradation. It is
 CC thought that AREs influence mRNA degradation rates. The invention relates
 CC to an in vitro system (consisting of a cell extract and an exogenous
 CC target RNA) that is able to recapitulate regulated RNA turnover of the
 CC target RNA. The system is used to identify agents that modulate
 CC stability, deadenylation or degradation of the target RNA, or endogenous
 CC molecules that participate in deadenylation/degradation of the target
 CC RNA. These modulators e.g AREs are preferably involved in cell growth and
 CC differentiation in mammals, especially where these processes are
 CC implicated in cell transformation and immune system dysfunction and are
 CC potential therapeutic agents, e.g. in conditions associated with abnormal
 CC expression of tumour necrosis factor-alpha. Examples of such diseases
 CC include sepsis, rheumatoid arthritis or inflammatory bowel disease. The
 CC system can also be used diagnostically to detect the molecular defects in
 CC such conditions and for development of improved gene delivery systems.
 XX
 SQ Sequence 33 BP; 10 A; 0 C; 0 G; 23 T; 0 other;
 Query Match 0.8%; Score 17; DB 21; Length 33;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

Qy 381 AATAAATAAATAATAA 397
Db 21 AATAAATAAATAATAA 5

RESULT 13
ABQ77985
ID ABQ77985 standard; DNA; 33 BP.
AC ABQ77985;
XX
XX 24-JAN-2003 (first entry)
DT
XX
DE DNA topoisomerase II 12.76 PCR primer, SEQ ID NO:6.
XX
KW DNA topoisomerase II 12.76; recombinant production; gene therapy;
KW malignant tumour; cancer; blood disease; HIV infection;
KW human immunodeficiency virus; immune disorder; inflammatory condition;
KW cytostatic; antiinflammatory; immunomodulator; PCR; primer; ss.
XX
OS Unidentified.
XX
XX CN1345979-A.
XX
XX 24-APR-2002.
XX
XX 26-SEP-2000; 2000CN-0125449.
XX
XX 26-SEP-2000; 2000CN-0125449.
XX
XX (SHAN-) SHANGHAI BLOWINDOW GENE DEV INC.
XX
XX Mao Y, Xie Y;
XX
XX WPI; 2002-575981/62.
XX
XX Novel polypeptide-DNA topoisomerase II 12.76 and polynucleotide for
XX encoding the polypeptide, useful for curing several diseases, such as
XX malignant tumor, hemopathy, HIV infection, immunological disease and
XX various inflammations -
XX
XX Example 4; Page 18 (Disclosure); 34pp; Chinese.
XX
CC The invention relates to DNA topoisomerase II 12.76 (ABB99905) and
CC nucleic acids encoding it (ABQ77974). The protein has a molecular weight
CC of 13 kD. The invention also relates to a method for the recombinant
CC production of the protein, an antagonist of the protein, and the use of
CC the protein, gene and antagonist in therapeutic applications. DNA
CC topoisomerase II 12.76 can be used in the treatment of a variety of
CC diseases such as malignant tumours, blood diseases, HIV (human
CC immunodeficiency virus) infection, immune disorders and inflammatory
CC conditions. Sequences ABQ77984-ABQ77985 represent PCR primers used in an
CC exemplification of the invention to amplify DNA topoisomerase II 12.76
CC cDNA for cloning.
XX
XX Sequence 33 BP; 9 A; 12 C; 6 G; 6 T; 0 other;
SQ
Query Match 0.8%; Score 17; DB 24; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1242 TTACAGGTGTAGCCAC 1258
Db 10 TTACAGGTGTAGCCAC 26

RESULT 14
AAA22698/c
ID AAA22698 standard; RNA; 17 BP.
XX
XX AAA22698;
AC
XX
XX 19-JUN-2000 (first entry)
DT

```

```

XX
DE
XX
KW Human; aryl hydrocarbon nuclear transport; ARNT; TIE-2; angiogenesis;
KW integrin alpha 6 subunit; integrin subunit beta 3; hairpin ribozyme;
KW hammerhead ribozyme; angiogenic factor; cyostatic; antidiabetic;
KW ophthalmologic; antiinflammatory; antiarthritic; antiposrotatic; ARMD;
KW dermatological; RNA cleavage; cancer; diabetic retinopathy; arthritis;
KW age related macular degeneration; inflammation; neovascular glaucoma;
KW myopic degeneration; psoriasis; verruca vulgaris; angiofibroma;
KW tuberosus sclerosis; pot-wine stain; Sturge Weber syndrome;
KW Kippel-Trenaunay-Weber syndrome; Osler-Weber-Rendu syndrome; ss.
XX
OS Homo sapiens.
XX
XX WO9950403-A2.
XX
XX 07-OCT-1999.
XX
XX 24-MAR-1999; 99WO-US06507.
XX
XX 27-MAR-1998; 98US-0079678.
XX
XX (RIBO-) RIBOZYME PHARM INC.
XX
XX Pavco PA, Roberts E, Jarvis T, Coeshott C, McSwiggen JA;
XX WPI; 1999-591315/50.
XX
XX Novel ribozymes for modulating the synthesis, expression and/or
XX stability of an mRNA encoding an angiogenic factors -
XX
XX Claim 54; Page 236; 305pp; English.
XX
CC The present invention describes enzymatic nucleic acid molecules with
CC RNA cleaving activity, which specifically cleave RNA encoded by an aryl
CC hydrocarbon nuclear transporter (ARNT) gene, an integrin subunit beta 3
CC gene, an integrin alpha 6 subunit gene, or a Tie-2 gene. AAA16775 to
CC AAA17167 and AAA17561 to AAA17622 represent ribozyme sequences for ARNT,
CC and AAA17168 to AAA17560 and AAA17623 to AAA17684 represent their
CC corresponding target sequences; AAA17685 to AAA18385 and AAA19087 to
CC AAA19154 represent ribozyme sequences for Tie-2, and AAA18386 to AAA19086
CC and AAA19155 to AAA19222 represent their corresponding target sequences;
CC AAA19223 to AAA20361 and AAA21501 to AAA21595 represent ribozyme
CC sequences for integrin alpha 6 subunit, and AAA20362 to AAA21500 and
CC AAA21596 to AAA21688 represent their corresponding target sequences;
CC AAA1689 to AAA22475 and AAA23263 to AAA23342 represent ribozyme sequences
CC for integrin subunit beta 3, and AAA22476 to AAA23262, AAA23343 to
CC AAA23422 represent their corresponding target sequences. The ribozymes of
CC the invention are used for modulating the synthesis, expression and/or
CC stability of an mRNA encoding angiogenic factor, especially ARNT,
CC integrin subunit beta-3, integrin subunit alpha-6, or Tie-2. They are
CC especially used to treat cancer, diabetic retinopathy, age related
CC macular degeneration (ARMD), inflammation, and arthritis, as well as
CC neovascular glaucoma, myopic degeneration, psoriasis, verruca vulgaris,
CC angiofibroma of tuberosus sclerosis, pot-wine stains, Sturge Weber
CC syndrome, Kippel-Trenaunay-Weber syndrome, Osler-Weber-Rendu syndrome,
CC and other syndromes and diseases related to the levels of ARNT, Tie-2,
CC integrin subunit alpha-6, or integrin subunit beta-3.
XX
XX Sequence 17 BP; 5 A; 0 C; 0 G; 12 U; 0 other;
SQ
Query Match 0.8%; Score 16; DB 20; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.4e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 380 AATAAATAAATAATAA 395
Db 16 AATAAATAAATAATAA 1

RESULT 15
AAA22708/c

```

ID AAA22708 standard; RNA; 17 BP.
XX AC
XX AAA22708;
XX DT 19-JUN-2000 (first entry)
XX DE Integrin subunit beta 3 substrate sequence SEQ ID NO:5934.
XX DE
XX Human; aryl hydrocarbon nuclear transport; ARNT; Tie-2; angiogenesis;
KW integrin alpha 6 subunit; integrin subunit beta 3; hairpin ribozyme;
KW hammerhead ribozyme; angiogenic factor; cytostatic; antidiabetic;
KW ophthalmologic; antiinflammatory; antiarthritic; antipsoriatic; ARMD;
KW dermatologic; RNA cleavage; cancer; diabetic retinopathy; arthritis;
KW age related macular degeneration; inflammation; neovascular glaucoma;
KW myopic degeneration; psoriasis; verruca vulgaris; angiofibroma;
KW tuberosus sclerosis; pot-wine stain; Sturge Weber syndrome; ss.
KW Kippel-Trenaunay-Weber syndrome; Osler-Weber-Rendu syndrome; ss.
XX OS Homo sapiens.
XX PN WO9950403-A2.
XX PD 07-OCT-1999.
XX PF 24-MAR-1999; 99WO-US06507.
XX PR 27-MAR-1998; 98US-0079678.
XX PA (RIBO-) RIBOZYME PHARM INC.
XX PI Pavco PA, Roberts E, Jarvis T, Coeshott C, McSwiggen JA;
XX WPI; 1999-591315/50.
XX PT Novel ribozymes for modulating the synthesis, expression and/or
XX stability of an mRNA encoding an angiogenic factors
XX Claim 54; Page 237; 305pp; English.
XX The present invention describes enzymatic nucleic acid molecules with
CC RNA cleaving activity, which specifically cleave RNA encoded by an aryl
CC hydrocarbon nuclear transporter (ARNT) gene, an integrin subunit beta 3
CC gene, an integrin alpha 6 subunit gene, or a Tie-2 gene. AAA16775 to
CC AAA17167 and AAA17561 to AAA17622 represent ribozyme sequences for ARNT,
CC and AAA17168 to AAA17560 and AAA17623 to AAA17684 represent their
CC corresponding target sequences; AAA17685 to AAA18385 and AAA19087 to
CC AAA19154 represent ribozyme sequences for Tie-2, and AAA18386 to AAA19086
CC and AAA19155 to AAA19222 represent their corresponding target sequences;
CC AAA19223 to AAA20361 and AAA21501 to AAA21595 represent ribozyme
CC sequences for integrin alpha 6 subunit, and AAA20362 to AAA21500 and
CC AAA21596 to AAA21688 represent their corresponding target sequences;
CC AAA21689 to AAA22475 and AAA23263 to AAA23342 represent ribozyme sequence
CC for integrin subunit beta 3, and AAA22476 to AAA23262, AAA23343 to
CC AAA23422 represent their corresponding target sequences. The ribozymes of
CC the invention are used for modulating the synthesis, expression and/or
CC stability of an mRNA encoding angiogenic factor, especially ARNT,
CC integrin subunit beta-3, integrin subunit alpha-6, or Tie-2. They are
CC especially used to treat cancer, diabetic retinopathy, age related
CC macular degeneration (ARMD), inflammation, and arthritis, as well as
CC neovascular glaucoma, myopic degeneration, psoriasis, verruca vulgaris,
CC angiofibroma of tuberosus sclerosis, pot-wine stains, Sturge Weber
CC syndrome, Kippel-Trenaunay-Weber syndrome, Osler-Weber-Rendu syndrome,
CC and other syndromes and diseases related to the levels of ARNT, Tie-2,
CC integrin subunit alpha-6, or integrin subunit beta-3.
XX Sequence 17 BP; 3 A; 0 C; 1 G; 13 U; 0 other;
XX Query Match 0.8%; Score 16; DB 20; Length 17;
XX Best Local Similarity 100.0%; Pred. No. 3.4e+03;
XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 378 AAAAATAAATAATAA 393
|||||

Db 16 AAAAATAAATAATAA 1

Search completed: November 14, 2003, 11:58:55
Job time : 537.347 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 14, 2003, 11:17:54 ; Search time 4313.19 Seconds
(without alignments)
11686.822 Million cell updates/sec

Title: US-10-005-337A-2
Perfect score: 2074
Sequence: 1 ctgcagcaagttacttaatg.....acaagactcttcagccaac 2074

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0
Total number of hits satisfying chosen parameters: 124404

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_ptg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18	0.9	28	14 H63106	H63106 yr48a04.s1
C 2	18	0.9	32	14 H70643	H70643 yu18h07.s1
C 3	18	0.9	38	9 AU264662	AU264662 AU264662
C 4	18	0.9	42	9 AL588371	AL588371 AL588371

5	18	0.9	49	9 AU268128	AU268128 AU268128
6	17	0.8	30	29 AL943661	AL943661 Arabidops
7	16	0.8	35	28 AZ823918	AZ823918 2M0098P11
8	16	0.8	42	28 AZ981274	AZ981274 2M0258G09
9	15	0.7	19	28 AZ333223	AZ333223 1M0062P08
10	15	0.7	21	28 AZ786126	AZ786126 1M0031B16
11	15	0.7	25	28 AZ345473	AZ345473 1M0080P08
12	15	0.7	27	14 R07762	R07762 yf15d04.s1
13	15	0.7	28	9 AA911003	AA911003 ok57f12.s
14	15	0.7	28	14 T65402	T65402 yc73d01.s1
15	15	0.7	30	8 AU266910	AU266910 AU266910
16	15	0.7	32	14 H21549	H21549 ym7a07.r1
17	15	0.7	34	28 AZ345619	AZ345619 1M0080E23
18	15	0.7	35	29 TAI94C08P	TAI94C08P T. brucei
19	15	0.7	36	14 N27542	N27542 y01d08.s1
20	15	0.7	38	14 H30464	H30464 y058a02.r1
21	15	0.7	39	9 AU266450	AU266450 AU266450
22	15	0.7	40	9 AU254489	AU254489 AU254489
23	15	0.7	42	28 AZ307978	AZ307978 1M0010C02
24	15	0.7	42	28 AZ876064	AZ876064 2M0190K19
25	15	0.7	45	29 TAI38F04P	TAI38F04P T. brucei
26	15	0.7	47	28 AZ488792	AZ488792 1M0319E17
27	15	0.7	47	28 AZ793185	AZ793185 2M0046P12
28	15	0.7	49	9 AI971870	AI971870 wv29e02.x
29	15	0.7	49	9 AW827212	AW827212 xn09f10.y
30	15	0.7	50	9 AA865303	AA865303 OG88a05.s
31	15	0.7	50	9 AU102270	AU102270 AU102270
32	15	0.7	50	9 AU104255	AU104255 AU104255
33	15	0.7	50	9 AU105701	AU105701 AU105701
34	15	0.7	50	9 AU105705	AU105705 AU105705
35	15	0.7	50	9 AU105707	AU105707 AU105707
36	15	0.7	50	9 AU105708	AU105708 AU105708
37	15	0.7	19	28 AZ787003	AZ787003 2M0032O13
38	14	0.7	19	28 AZ799396	AZ799396 2M0056N18
39	14	0.7	20	28 AZ346143	AZ346143 1M0081P11
40	14	0.7	22	9 AW247689	AW247689 2820007.s
41	14	0.7	23	28 AZ939702	AZ939702 2M0198K15
42	14	0.7	24	28 AZ336547	AZ336547 1M0066E23
43	14	0.7	25	14 T80419	T80419 yd17d11.s1
44	14	0.7	25	14 T80419	T80419 yd17d11.s1
45	14	0.7	25	28 AZ331633	AZ331633 1M0059N11

ALIGNMENTS

RESULT 1
LOCUS H63106/c
DEFINITION yr48a04.s1 Soares fetal liver EST 11-OCT-1995
IMAGE:208494 3' similar to gb|U87917|HUMALNE441 Human carcinoma cell-derived Alu RNA transcript. (rRNA); gb:M91159 !!! ALU CLASS E
WARNING ENTRY !!!! (HUMAN);, mRNA sequence.

ACCESSION H63106
VERSION H63106.1 GI:1017907
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 28)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished
COMMENT Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

Insert Size: 3194
 High quality sequence starts: 1
 High quality sequence stops: 1
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Trace considered overall poor quality
 Insert Length: 3194 Std Error: 0.00
 Seq primer: Promega -21ml3
 High quality sequence stop: 1.
 Location/Qualifiers
 1. .28

FEATURES

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3777625"
/db_xref="taxon:9606"
/clone="IMAGE:208494"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: p7T3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5'- AACTGGAAGATTAAATTAAGATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Patima Bonaldo."

```

6 a

9 c 5 g

8 t

0;

Gaps

0;

Indels

0;

Mismatches

0;

Conservative

0;

Pred. No.

1.2e+04;

Length

28;

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 0.9%; Score 18; DB 14; Length 28;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1237 TGGGATTACAGGTGAAG 1254

|||||

Db 18 TGGGATTACAGGTGAAG 1

RESULT 2

H70643/c

LOCUS

DEFINITION

H70643 yul8h07.s1 Soares fetal liver spleen INFLS Homo sapiens EST 26-OCT-1995
 IMAGE:234205 3' similar to gb:D10202 PLATELET ACTIVATING FACTOR
 RECEPTOR (HUMAN) ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

COMMENT

H70643 32 bp mRNA linear EST 26-OCT-1995
 yul8h07.s1 Soares fetal liver spleen INFLS Homo sapiens EST 26-OCT-1995
 IMAGE:234205 3' similar to gb:D10202 PLATELET ACTIVATING FACTOR
 RECEPTOR (HUMAN) ; mRNA sequence.

Unpublished

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

Insert Size: 1926

High quality sequence starts: 1

High quality sequence stops: 1

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.
 Trace considered overall poor quality
 Insert Length: 1926 Std Error: 0.00
 Seq primer: Promega -21ml3
 High quality sequence stop: 1.
 Location/Qualifiers
 1. .32

FEATURES

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3787382"
/db_xref="taxon:9606"
/clone="IMAGE:234205"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: p7T3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5'- AACTGGAAGATTAAATTAAGATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Patima Bonaldo."

```

7 a

9 c 7 g

9 t

0;

Gaps

0;

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 0.9%; Score 18; DB 14; Length 32;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1237 TGGGATTACAGGTGAAG 1254

|||||

Db 22 TGGGATTACAGGTGAAG 5

RESULT 3

AU264662/c

LOCUS

DEFINITION

AU264662 VS Dictyostelium discoideum cDNA clone VSD856 5', mRNA
 sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AU264662 38 bp mRNA linear EST 10-MAY-2002
 AU264662 VS Dictyostelium discoideum cDNA clone VSD856 5', mRNA
 sequence.

Unpublished

Contact: Hideko Urushihara

Institute of Biological Sciences

University of Tsukuba

1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan

Tel: 81-298-53-4664

Fax: 81-298-53-6614

Email: hideko@biol.tsukuba.ac.jp.

Location/Qualifiers

1. .38

/organism="Dictyostelium discoideum"

/mol_type="mRNA"

/strain="AX4"

/db_xref="taxon:44689"

/clone="VSD856"

/sex="mat A"

/dev_stage="vegetative"

/clone_lib="VS"

25 a

2 c 1 g 10 t

```

Query Match      0.9%; Score 18; DB 9; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1130 ATATATATATATAGTATT 1147
    |||||
Db 18 ATATATATATATAGTATT 1

RESULT 4
AL588371
LOCUS      42 bp mRNA linear EST 02-MAR-2001
DEFINITION ROS071D10, mRNA sequence.
ACCESSION  AL588371
VERSION     AL588371.1 GI:13193405
KEYWORDS   EST.
SOURCE     Gallus gallus (chicken)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 42)
AUTHORS   Murray, F.
TITLE     BP Chicken Brain Library
JOURNAL   Unpublished
COMMENT   Contact: Frazer Murray
            Dept. Genomics and Bioinformatics
            Roslin Institute
            Roslin, Midlothian, EH25 9PS, UK
            Tel: +44 (0)131 527 4200
            Fax: +44 (0)131 440 0434
            Email: frazer.murray@bbsrc.ac.uk
            GCGGCCGCTTTT TTTT TTTT TTTT 3' Poly A RNA purchased from Clontech
            (*6854-

Seq primer: 77.
            Location/Qualifiers
            1..42
            /organism="Gallus gallus"
            /mol_type="mRNA"
            /db_xref="taxon:9031"
            /clone="ROS071D10"
            /issue_type="Brain"
            /dev_stage="Unknown"
            /lab_host="DH10B"
            /clone_lib="BP Chicken Brain Library"
            /notes="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; Cloned
            unidirectionally. Primer: Oligo dT. 5' adaptor sequence:
            5' TCGACCTCGAG 3'; 3' adaptor sequence: 5'
            GCGGCCGCTTTT TTTT TTTT TTTT 3' Poly A RNA purchased from
            Clontech (*6854-1)"
            19 a 5 c 8 g 10 t

BASE COUNT      19 a 5 c 8 g 10 t
ORIGIN

Query Match      0.9%; Score 18; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 ATTAATAATAATAATAA 392
    |||||
Db 6 ATTAATAATAATAATAA 23

RESULT 5
AU268128
LOCUS      49 bp mRNA linear EST 10-MAY-2002
DEFINITION AU268128 VS Dictyostelium discoideum cDNA clone VSH886 5', mRNA
            sequence.
ACCESSION  AU268128
VERSION     AU268128.1 GI:20526926
KEYWORDS   EST.
SOURCE     Dictyostelium discoideum

```

```

ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
AUTHORS 1 (bases 1 to 49)
        Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
        Takeuchi,I., Kohara,Y. and Tanaka,Y.
TITLE    Population analysis of cDNAs from unicellular and multicellular
        stages of Dictyostelium discoideum
JOURNAL  Unpublished
COMMENT  Contact: Hideko Urushihara
        Institute of Biological Sciences
        University of Tsukuba
        1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
        Tel: 81-298-53-4664
        Fax: 81-298-53-6614
        Email: hideko@biol.tsukuba.ac.jp.

FEATURES
        Location/Qualifiers
        1..49
        /organism="Dictyostelium discoideum"
        /mol_type="mRNA"
        /strains="AX4"
        /db_xref="taxon:44689"
        /clone="VSH886"
        /sex="mat A"
        /dev_stage="vegetative"
        /clone_lib="VS"

BASE COUNT      35 a 2 c 1 g 10 t 1 others
ORIGIN

Query Match      0.9%; Score 18; DB 9; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 AAATAATAATAATAATAA 397
    |||||
Db 10 AAATAATAATAATAATAA 27

RESULT 6
AL943661/c
LOCUS      30 bp DNA linear GSS 24-OCT-2002
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-276A12-015154,
        Genomic survey sequence.
ACCESSION  AL943661
VERSION     AL943661.1 GI:24400267
KEYWORDS   GSS.
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
        Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
        rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE  1
AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoveer,P., Dekker,K., Saedler,H.
        and Weissshaar,B.
TITLE     A pipeline for automated high-throughput generation of FSTs
        (flanking sequence tags) from Arabidopsis thaliana T-DNA
        transformed lines
JOURNAL  Unpublished
REFERENCE  2
AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weissshaar,B.
TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
        for flanking sequence tag based reverse genetics
JOURNAL  Unpublished
REFERENCE  3 (bases 1 to 30)
AUTHORS   Li,Y., Rosso,M., Strizhov,N. and Weissshaar,B.
TITLE     Direct Submission
JOURNAL  Submitted (21-OCT-2002) Weissshaar B., Max-Planck-Institut fuer
        Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
        This sequence is recovered from the left border of the T-DNA. It
        indicates an insertion within the locus defined by clone MCK7. The
        sequences are generated at the MPI for Plant Breeding Research in
        the context of the GABI-Kat project. GABI-Kat is part of the German
        Plant Genomics program designated 'GABI'. Information on line
        availability can be found at:

```

http://www.mpiz-koeln.mpg.de/GABI-Kat/.

FEATURES

Location/Qualifiers
1..30
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="CK-276A12-015154"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector PAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

BASE COUNT 15 a 1 c 0 g 14 t
ORIGIN

Query Match 0.8%; Score 17; DB 29; Length 30;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1130 ATATATATATAGTAT 1146
|||||
DB 17 ATATATATATAGTAT 1

RESULT 7

AZ823918 35 bp DNA linear GSS 20-FEB-2001
LOCUS 2M0098P11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC2M0098P11 F, genomic survey sequence.
ACCESSION AZ823918
VERSION AZ823918.1 GI:12993826
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 35)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0098 row: P column: 11
Seq primer: CGTTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 35.
Location/Qualifiers
1..35
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0098P11"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"

FEATURES

Location/Qualifiers
1..35
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0098P11"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 16 a 1 c 2 g 16 t
ORIGIN

Query Match 0.8%; Score 16; DB 28; Length 35;
Best Local Similarity 100.0%; Pred. No. 8.7e+04;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1131 TATATATATAGTAT 1146
|||||
DB 10 TATATATATAGTAT 25

RESULT 8

AZ981274 42 bp DNA linear GSS 27-APR-2001
LOCUS 2M0258G09R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
DEFINITION clone UUGC2M0258G09 R, genomic survey sequence.
ACCESSION AZ981274
VERSION AZ981274.1 GI:13852501
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 42)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0258 row: G column: 09
Seq primer: CACACGGAACAGGTATGACC
Class: plasmid ends
High quality sequence stop: 42.
Location/Qualifiers
1..42
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0258G09"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC2M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (female); Purified genomic DNA from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (GI|4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 16 a 7 c 9 g 10 t
 ORIGIN
 Query Match 0.8%; Score 16; DB 28; Length 42;
 Best Local Similarity 100.0%; Pred. No. 8.8e+04;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1323 AAAAATGCTCCCAATTA 1338
 |||||
 DB 18 AAAAATGCTCCCAATTA 33

RESULT 9
 AZ333223/c
 LOCUS
 DEFINITION 19 bp DNA linear GSS 29-SEP-2000
 1M0062P08F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0062P08 F, genomic survey sequence.

ACCESSION AZ333223
 VERSION AZ333223.1 GI:10397629
 KEYWORDS
 SOURCE GSS.

ORGANISM Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 19)
 REFERENCE
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
 and Wright,D.,Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

JOURNAL
 COMMENT Unpublished
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0062 row: F column: 08
 Seq primer: CGTTGTAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 19.

FEATURES
 Location/Qualifiers
 1..19

/organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0062P08"
 /sex="Male"

FEATURES
 source

1..21
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0031B16"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male); Purified genomic DNA from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (GI|4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 5 a 1 c 0 g 13 t
 ORIGIN

Query Match 0.7%; Score 15; DB 28; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2.3e+05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 AAAATAAATAATAA 393
 |||||
 DB 16 AAAATAAATAATAA 2

RESULT 10
 AZ786126/c

LOCUS
 DEFINITION 21 bp DNA linear GSS 16-FEB-2001
 1M0031B16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0031B16 F, genomic survey sequence.

ACCESSION AZ786126
 VERSION AZ786126.1 GI:12923574
 KEYWORDS
 SOURCE GSS.

ORGANISM Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 21)
 REFERENCE
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
 and Wright,D.,Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

JOURNAL
 COMMENT Unpublished
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0031 row: B column: 16
 Seq primer: CGTTGTAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 21.

Location/Qualifiers

1..21
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0031B16"

```

/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      8 a      2 c      4 g      7 t
ORIGIN
Query Match      0.7%; Score 15; DB 28; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.3e+05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1128 TCATATATATATATA 1142
      |||||
DB 15 TCATATATATATATA 1

```

```

RESULT 11
AZ345473
LOCUS
DEFINITION
1M0080P08F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0080P08 F, genomic survey sequence.
ACCESSION
AZ345473
VERSION
AZ345473.1 GI:10424710
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 25)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0080 row: P column: 08
Seq primer: CGTTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 25.
Location/Qualifiers
1..25
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

```

```

FEATURES
source

```

```

/clone="UUGC1M0080P08"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      14 a      3 c      1 g      7 t
ORIGIN
Query Match      0.7%; Score 15; DB 28; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.3e+05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 379 AAATAATAATAATAA 393
      |||||
DB 10 AAATAATAATAATAA 24

```

```

RESULT 12
R07762
LOCUS
DEFINITION
Yf15d04.s1 Soares fetal liver spleen INFLS Homo sapiens CDNA clone
IMAGE:126919 3, similar to gb:M81181 SODIUM/POTASSIUM-TRANSPORTING
ATPASE BETA-2 (HUMAN); mRNA sequence.
ACCESSION
R07762
VERSION
R07762.1 GI:759685
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 27)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 964
High quality sequence starts: 1 High quality sequence stops: 1
Source: IMAGE Consortium, LLNL This clone is available royalty-free
through LLNL; contact the IMAGE Consortium (info@image.llnl.gov)
for further information. Trace considered overall poor quality
Insert Length: 964 Std Error: 0.00
Seq primer: -21m13
High quality sequence stop: 1.
Location/Qualifiers
1..27
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

FEATURES
source

```



```

/db_xref="CDB:479080"
/db_xref="taxon:9606"
/clone="IMAGE:126919"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/notes="Organ: Liver and Spleen; Vector: pVT3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5'-AACTGGAGAATAAATAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pVT3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      7 a      4 c      9 g      6 t
ORIGIN
Query Match      0.7%; Score 15; DB 14; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1234 TGTGGGATTACAGG 1248
Db      |||||||
5 TGTGGGATTACAGG 19

RESULT 13
AA911003
LOCUS
DEFINITION
OK57112.s1 NCI CGAP Lei2 Homo sapiens cDNA clone IMAGE:1518095 3'
similar to TR:Q41805 Q41805 EXTENSIN-LIKE PROTEIN PRECURSOR.
;contains element MSR1 repetitive element ;, mRNA sequence.
ACCESSION      AA911003
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 28)
REFERENCE
AUTHORS
Rifkin,L., Rohlfsing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-f@mail.nih.gov
unknown library type
Trace considered overall poor quality
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
FEATURES
source
1..28
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1518095"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lei2"
/notes="Organ: soft tissue; Vector: pVT3D-Pac (Pharmacia)
with a modified polylinker; Site 1: Not I; Site 2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'-AACTGGAGAATTCGGCCGCAATCGTTTTTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pVT3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      9 a      13 c      3 g      3 t
ORIGIN

```

```

Query Match      0.7%; Score 15; DB 9; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.3e+05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 797 ATAAAAATACCCCA 811
Db      |||||||
5 ATAAAAATACCCCA 19

RESULT 14
T65402
LOCUS
DEFINITION
yc73d01.s1 Soares infant brain IN1B Homo sapiens cDNA clone
IMAGE:21732 3' similar to gb|M87917|HUMANLE441 Human carcinoma
cell-derived Alu RNA transcript. (rRNA); gb:J04513 HEPARIN-BINDING
GROWTH FACTOR PRECURSOR 2 (HUMAN);, mRNA sequence.
ACCESSION      T65402
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 28)
REFERENCE
AUTHORS
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfsing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
JOURNAL
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert Size: 3006
High quality sequence starts: 1 High quality sequence stops: 1
Source: IMAGE Consortium, LLNL This clone is available royalty-free
through LLNL; contact the IMAGE Consortium (info@image.llnl.gov)
for further information. Trace considered overall poor quality
Insert Length: 3006 Std Error: 0.00
Seq primer: -21ml3
High quality sequence stop: 1.
FEATURES
source
1..28
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:394079"
/db_xref="taxon:9606"
/clone="IMAGE:21732"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain IN1B"
/notes="Organ: whole brain; Vector: Lafmid BA; Site 1: Not
I; Site 2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5',
AACTGGAGAATTCGGCCGCAATTTTTTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lafmid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      7 a      4 c      10 g      7 t
ORIGIN
Query Match      0.7%; Score 15; DB 14; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.3e+05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1237 TGGGATTACAGTGT 1251
|||||||

```

Db 10 TGGGATTACAGGTGT 24

RESULT 15
AZ345640/c
LOCUS
DEFINITION
AZ345640 28 bp DNA linear GSS 29-SEP-2000
1M0080124F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0080124 F, genomic survey sequence.

ACCESSION
AZ345640
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)

REFERENCE
AUTHORS
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
.M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 306, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0080 row: 1 column: 24
Seq primer: CGTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 28.

FEATURES
source
1..28
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0080124"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 6 a 0 c 0 g 22 t
ORIGIN

Query Match 0.7%; Score 15; DB 28; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.3e+05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 AAAATAAATAATAA 393

Db 28 AAAATAAATAATAA 14

Search completed: November 14, 2003, 19:05:52
Job time : 4320.19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 06:49:07 ; Search time 117.926 Seconds
(without alignments)
7762.738 Million cell updates/sec

Title: US-10-005-337A-2

Perfect score: 2074

Sequence: 1 ctgcagaagttacttaatg.....acaagactcttcagccaac 2074

Scoring table:

OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 744296

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18	0.9	34	1 US-08-434-503-29	Sequence 29, Appl
C 2	17	0.8	30	3 US-09-176-862-18	Sequence 18, Appl
C 3	16	0.8	19	3 US-09-280-409-44	Sequence 44, Appl
C 4	16	0.8	18	4 US-09-422-978-4951	Sequence 4951, Ap
C 5	16	0.8	46	4 US-09-641-638-1292	Sequence 1292, Ap
C 6	16	0.8	47	4 US-09-422-978-2343	Sequence 2343, Ap
C 7	16	0.8	47	4 US-09-422-978-2541	Sequence 2541, Ap
C 8	15	0.7	15	1 US-08-311-486C-191	Sequence 191, App
C 9	15	0.7	15	1 US-08-311-486C-192	Sequence 192, App
C 10	15	0.7	15	1 US-08-311-486C-199	Sequence 199, App
C 11	15	0.7	15	1 US-08-311-486C-200	Sequence 200, App
C 12	15	0.7	15	1 US-08-311-486C-201	Sequence 201, App
C 13	15	0.7	15	1 US-08-311-486C-712	Sequence 712, App
C 14	15	0.7	15	1 US-08-311-486C-713	Sequence 713, App
C 15	15	0.7	15	1 US-08-311-486C-714	Sequence 714, App
C 16	15	0.7	15	1 US-08-311-486C-715	Sequence 715, App
C 17	15	0.7	15	1 US-08-311-486C-716	Sequence 716, App
C 18	15	0.7	15	1 US-08-311-486C-721	Sequence 721, App
C 19	15	0.7	20	1 US-08-480-784-3	Sequence 3, Appli
C 20	15	0.7	20	1 US-08-483-553-3	Sequence 3, Appli
C 21	15	0.7	20	1 US-08-487-002-3	Sequence 3, Appli
C 22	15	0.7	20	1 US-08-483-554B-3	Sequence 3, Appli
C 23	15	0.7	20	1 US-08-488-011B-3	Sequence 3, Appli
C 24	15	0.7	20	3 US-08-850-727-3	Sequence 3, Appli
C 25	15	0.7	20	4 US-09-705-299-58	Sequence 58, Appl
C 26	15	0.7	20	4 US-09-791-211-75	Sequence 75, Appl
C 27	15	0.7	20	4 US-09-918-686-77	Sequence 77, Appl

28	15	0.7	20	4	US-09-679-299A-69	Sequence 69, Appl
C 29	15	0.7	20	5	PCT-US95-10202-3	Sequence 3, Appli
C 30	15	0.7	20	5	PCT-US95-10203-3	Sequence 3, Appli
C 31	15	0.7	20	5	PCT-US95-10220-3	Sequence 3, Appli
C 32	15	0.7	21	4	US-09-422-978-10125	Sequence 10125, A
C 33	15	0.7	28	2	US-08-859-998-147	Sequence 147, App
C 34	15	0.7	28	4	US-09-225-928-147	Sequence 147, App
C 35	15	0.7	28	4	US-09-225-201B-147	Sequence 147, App
C 36	15	0.7	47	4	US-09-641-638-1027	Sequence 1027, Ap
C 37	15	0.7	47	4	US-09-422-978-488	Sequence 488, App
C 38	15	0.7	47	4	US-09-422-978-2669	Sequence 2669, Ap
C 39	15	0.7	47	4	US-09-422-978-3399	Sequence 3399, Ap
C 40	15	0.7	48	1	US-08-171-389-18	Sequence 18, Appl
C 41	15	0.7	48	1	US-08-123-936-18	Sequence 18, Appl
C 42	15	0.7	48	2	US-08-475-228A-18	Sequence 18, Appl
C 43	15	0.7	48	3	US-08-482-080A-18	Sequence 18, Appl
C 44	15	0.7	48	4	US-09-354-947-18	Sequence 18, Appl
C 45	15	0.7	48	5	PCT-US93-12388-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-434-503-29/c
; Sequence 29, Application US/08434503
; Patent No. 5616490
; GENERAL INFORMATION:
; APPLICANT: Sean M. Sullivan
; APPLICANT: Kenneth G. Draper
; TITLE OF INVENTION: METHOD AND REAGENT FOR
; TITLE OF INVENTION: TREATMENT OF INFLAMMATORY
; TITLE OF INVENTION: DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 611 West Sixth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
; SOFTWARE: Wordperfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,503
; FILING DATE: 04-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,895
; FILING DATE: 19-JAN-1993
; APPLICATION NUMBER: 07/989,849
; FILING DATE: December 7, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 200/276
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-434-503-29

Query Match 0.9%; Score 18; DB 1; Length 34;
Best Local Similarity 100.0%; Pred. No. 60;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 AATAAATAAATAATAAA 397
|||||
Db 33 AATAAATAAATAATAAA 16

RESULT 2
US-09-176-862-18
; Sequence 18, Application US/09176862B
; Patent No. 6046319
; GENERAL INFORMATION:
; APPLICANT: Power, Christopher
; APPLICANT: Mayne, Michael B.
; TITLE OF INVENTION: ANTISENSE OLIGODEOXYNUCLEOTIDES REGULATING EXPRESSION
; FILE REFERENCE: 3045.00002
; CURRENT APPLICATION NUMBER: US/09/176,862B
; CURRENT FILING DATE: 1998-10-22
; EARLIER APPLICATION NUMBER: 60/062,718
; EARLIER FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
US-09-176-862-18

Query Match 0.8%; Score 17; DB 3; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 AATAAATAAATAATAAA 397
|||||
Db 4 AATAAATAAATAATAAA 20

RESULT 3
US-09-280-409-44
; Sequence 44, Application US/09280409
; Patent No. 6107092
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowsett
; APPLICANT: C. Frank Bennett
; APPLICANT: Bert W. O'Malley
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRA EXPRESSION
; FILE REFERENCE: RTS-0048
; CURRENT APPLICATION NUMBER: US/09/280,409
; CURRENT FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 146
; SEQ ID NO 44
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-280-409-44

Query Match 0.8%; Score 16; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 601 ACCATCCACTGACTGA 616
|||||
Db 1 ACCATCCACTGACTGA 16

RESULT 4
US-09-422-978-4951
; Sequence 4951, Application US/09422978

; Patent No. 6537751
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density....
; FILE REFERENCE: GENSET.020CPI
; CURRENT APPLICATION NUMBER: US/09/422,978
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/298,850
; EARLIER FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,732
; EARLIER FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: US 60/082,614
; EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 4951
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1..19
; OTHER INFORMATION: upstream amplification primer 99-18974 for SEQ 1017,
US-09-422-978-4951

Query Match 0.8%; Score 16; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2046 GAAAAACATACAGAC 2061
|||||
Db 2 GAAAAACATACAGAC 17

RESULT 5
US-09-641-638-1292/c
; Sequence 1292, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 1292
; LENGTH: 46
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 21
; OTHER INFORMATION: 10-510-173 : variable motif ATTTA or TTTTTT
US-09-641-638-1292

Query Match 0.8%; Score 16; DB 4; Length 46;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 AAAAATAAATAATAA 393
Db 25 AAAAATAAATAATAA 10

RESULT 6

US-09-422-978-2343
; Sequence 2343, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET.020CPI
; CURRENT APPLICATION NUMBER: US/09/422,978
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/298,850
; EARLIER FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,732
; EARLIER FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: US 60/082,614
; EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 2343
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-10521-296 : polymorphic base C or T
US-09-422-978-2343

Query Match 0.8%; Score 16; DB 4; Length 47;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 AAAAATAAATAATAA 393
Db 6 AAAAATAAATAATAA 21

RESULT 7

US-09-422-978-2541/c
; Sequence 2541, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET.020CPI
; CURRENT APPLICATION NUMBER: US/09/422,978
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/298,850
; EARLIER FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,732
; EARLIER FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: US 60/082,614
; EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 2541
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-11685-200 : polymorphic base T or C
US-09-422-978-2541

Query Match 0.8%; Score 16; DB 4; Length 47;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1253 AGCCACTGCACCGGC 1268
Db 21 AGCCACTGCACCGGC 6

RESULT 8

US-08-311-486C-191/c
; Sequence 191, Application US/08311486C
; Patent No. 5811300
; GENERAL INFORMATION:
; APPLICANT: Sean Sullivan
; APPLICANT: Kenneth Draper
; APPLICANT: Kevin Kisich
; APPLICANT: Dan T. Stinchcomb
; APPLICANT: James McSwiggen
; TITLE OF INVENTION: RIBOZYME TREATMENT OF
; TITLE OF INVENTION: DISEASES OR CONDITIONS
; TITLE OF INVENTION: RELATED TO LEVELS OF
; TITLE OF INVENTION: TNF-
; NUMBER OF SEQUENCES: 1157
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,486C
; FILING DATE: September 23, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/008,895
; FILING DATE: January 19, 1993
; APPLICATION NUMBER: 07/989,849
; FILING DATE: December 7, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 209/166
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 191:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-311-486C-191

Query Match 0.7%; Score 15; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 ATAAATAAATAATAA 396
Db 15 ATAAATAAATAATAA 1

RESULT 9
 US-08-311-486C-192/c
 ; Sequence 192, Application US/08311486C
 ; Patent No. 5811300
 ; GENERAL INFORMATION:
 ; APPLICANT: Sean Sullivan
 ; APPLICANT: Kenneth Draper
 ; APPLICANT: Kevin Kisich
 ; APPLICANT: Dan T. Stinchcomb
 ; APPLICANT: James McSwiggen
 ; TITLE OF INVENTION: RIBOZYME TREATMENT OF
 ; DISEASES OR CONDITIONS
 ; TITLE OF INVENTION: RELATED TO LEVELS OF
 ; TITLE OF INVENTION: TNF-
 ; NUMBER OF SEQUENCES: 1157
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: 633 West Fifth Street
 ; CITY: Suite 4700
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: U.S.A.
 ; ZIP: 90071-2066
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; MEDIUM TYPE: storage
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0
 ; SOFTWARE: Word Perfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/311,486C
 ; FILING DATE: September 23, 1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA: including application
 ; PRIOR APPLICATION DATA: described below:
 ; APPLICATION NUMBER: 08/008,895
 ; FILING DATE: January 19, 1993
 ; APPLICATION NUMBER: 07/989,849
 ; FILING DATE: December 7, 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Warburg, Richard J.
 ; REGISTRATION NUMBER: 32,327
 ; REFERENCE/DOCKET NUMBER: 209/166
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (213) 489-1600
 ; TELEFAX: (213) 955-0440
 ; TELEX: 67-3510
 ; INFORMATION FOR SEQ ID NO: 192:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 15 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-311-486C-192

Query Match 0.7%; Score 15; DB 1; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 AATAAATAAATA 395
 DB 15 AATAAATAAATA 1

RESULT 10
 US-08-311-486C-199/c
 ; Sequence 199, Application US/08311486C
 ; Patent No. 5811300
 ; GENERAL INFORMATION:
 ; APPLICANT: Sean Sullivan
 ; APPLICANT: Kenneth Draper
 ; APPLICANT: Kevin Kisich

; APPLICANT: Dan T. Stinchcomb
 ; APPLICANT: James McSwiggen
 ; TITLE OF INVENTION: RIBOZYME TREATMENT OF
 ; DISEASES OR CONDITIONS
 ; TITLE OF INVENTION: RELATED TO LEVELS OF
 ; TITLE OF INVENTION: TNF-
 ; NUMBER OF SEQUENCES: 1157
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: 633 West Fifth Street
 ; CITY: Suite 4700
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: U.S.A.
 ; ZIP: 90071-2066
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; MEDIUM TYPE: storage
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0
 ; SOFTWARE: Word Perfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/311,486C
 ; FILING DATE: September 23, 1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA: including application
 ; PRIOR APPLICATION DATA: described below:
 ; APPLICATION NUMBER: 08/008,895
 ; FILING DATE: January 19, 1993
 ; APPLICATION NUMBER: 07/989,849
 ; FILING DATE: December 7, 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Warburg, Richard J.
 ; REGISTRATION NUMBER: 32,327
 ; REFERENCE/DOCKET NUMBER: 209/166
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (213) 489-1600
 ; TELEFAX: (213) 955-0440
 ; TELEX: 67-3510
 ; INFORMATION FOR SEQ ID NO: 199:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 15 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-311-486C-199

Query Match 0.7%; Score 15; DB 1; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 ATAAATAAATAATA 396
 DB 15 ATAAATAAATAATA 1

RESULT 11
 US-08-311-486C-200/c
 ; Sequence 200, Application US/08311486C
 ; Patent No. 5811300
 ; GENERAL INFORMATION:
 ; APPLICANT: Sean Sullivan
 ; APPLICANT: Kenneth Draper
 ; APPLICANT: Kevin Kisich
 ; APPLICANT: Dan T. Stinchcomb
 ; APPLICANT: James McSwiggen
 ; TITLE OF INVENTION: RIBOZYME TREATMENT OF
 ; DISEASES OR CONDITIONS
 ; TITLE OF INVENTION: RELATED TO LEVELS OF
 ; TITLE OF INVENTION: TNF-
 ; NUMBER OF SEQUENCES: 1157
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,486C
FILING DATE: September 23, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/008,895
FILING DATE: January 19, 1993
APPLICATION NUMBER: 07/989,849
FILING DATE: December 7, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/166
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 200:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-311-486C-200

two

Query Match 0.7%; Score 15; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 381 AATAAATAATAATA 395
Db 15 AATAAATAATAATA 1
RESULT 12
US-08-311-486C-201/c
Sequence 201, Application US/08311486C
Patent No. 5811300
GENERAL INFORMATION:
APPLICANT: Sean Sullivan
APPLICANT: Kenneth Draper
APPLICANT: Kevin Kisich
APPLICANT: Dan T. Stinchcomb
APPLICANT: James McSwiggan
TITLE OF INVENTION: RIBOZYME TREATMENT OF
TITLE OF INVENTION: DISEASES OR CONDITIONS
TITLE OF INVENTION: RELATED TO LEVELS OF
TITLE OF INVENTION: TNF-
NUMBER OF SEQUENCES: 1157
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,486C
FILING DATE: September 23, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/008,895
FILING DATE: January 19, 1993
APPLICATION NUMBER: 07/989,849
FILING DATE: December 7, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/166
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 201:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-311-486C-201

two

Query Match 0.7%; Score 15; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 380 AATAAATAATAATA 394
Db 15 AATAAATAATAATA 1

RESULT 13
US-08-311-486C-712/c
Sequence 712, Application US/08311486C
Patent No. 5811300
GENERAL INFORMATION:
APPLICANT: Sean Sullivan
APPLICANT: Kenneth Draper
APPLICANT: Kevin Kisich
APPLICANT: Dan T. Stinchcomb
APPLICANT: James McSwiggan
TITLE OF INVENTION: RIBOZYME TREATMENT OF
TITLE OF INVENTION: DISEASES OR CONDITIONS
TITLE OF INVENTION: RELATED TO LEVELS OF
TITLE OF INVENTION: TNF-
NUMBER OF SEQUENCES: 1157
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,486C
FILING DATE: September 23, 1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/008,895
FILING DATE: January 19, 1993
APPLICATION NUMBER: 07/989,849
FILING DATE: December 7, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/166
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 712:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-311-486C-712

Query Match 0.7%; Score 15; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. NO. 1.7e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 ATAAATAAATAATAA 396
DB 15 ATAAATAAATAATAA 1

RESULT 14
US-08-311-486C-713/c
Sequence 713, Application US/08311486C
Patent No. 5811300
GENERAL INFORMATION:

APPLICANT: Sean Sullivan
APPLICANT: Kenneth Draper
APPLICANT: Kevin Kisich
APPLICANT: Dan T. Stinchcomb
APPLICANT: James McSwigen
TITLE OF INVENTION: RIBOZYME TREATMENT OF
TITLE OF INVENTION: DISEASES OR CONDITIONS
TITLE OF INVENTION: RELATED TO LEVELS OF
TITLE OF INVENTION: TNF-
NUMBER OF SEQUENCES: 1157
CORRESPONDENCE ADDRESS:

ADDRESSER: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,486C
FILING DATE: September 23, 1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/008,895
FILING DATE: January 19, 1993
APPLICATION NUMBER: 07/989,849
FILING DATE: December 7, 1992

two

ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/166
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 713:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-311-486C-713

Query Match 0.7%; Score 15; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. NO. 1.7e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 ATAAATAAATAATAA 396
DB 15 ATAAATAAATAATAA 1

RESULT 15
US-08-311-486C-714/c
Sequence 714, Application US/08311486C
Patent No. 5811300
GENERAL INFORMATION:

APPLICANT: Sean Sullivan
APPLICANT: Kenneth Draper
APPLICANT: Kevin Kisich
APPLICANT: Dan T. Stinchcomb
APPLICANT: James McSwigen
TITLE OF INVENTION: RIBOZYME TREATMENT OF
TITLE OF INVENTION: DISEASES OR CONDITIONS
TITLE OF INVENTION: RELATED TO LEVELS OF
TITLE OF INVENTION: TNF-
NUMBER OF SEQUENCES: 1157
CORRESPONDENCE ADDRESS:

ADDRESSER: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,486C
FILING DATE: September 23, 1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/008,895
FILING DATE: January 19, 1993
APPLICATION NUMBER: 07/989,849
FILING DATE: December 7, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/166
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

two

; INFORMATION FOR SEQ ID NO: 714:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 15 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-311-486C-714

Query Match

Best Local Similarity 0.7%; Score 15; DB 1; Length 15;

Matches 15; Conservativity 100.0%; Pred. No. 1.7e+03;

Mismatches 0; Indels 0; Gaps 0;

Qy 382 ATAAATAAATAATAA 396

Db 15 ATAAATAAATAATAA 1

Search completed: November 14, 2003, 16:31:58

Job time : 123.926 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 11:39:49 ; Search time 626.131 Seconds
(without alignments)
10825.622 Million cell updates/sec

Title: US-10-005-337A-2

Perfect score: 2074

Sequence: 1 ctgcagcaagtacttaatg.....acaagactcttcagccaac 2074

Scoring table:

OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 2169961 seqs, 1634102185 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1195920

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	1.0	30	14	US-10-085-906-6
C 2	18	0.9	19	12	US-10-224-005-6
C 3	18	0.9	19	12	US-10-224-005-6
C 4	18	0.9	34	10	US-09-955-462A-2
C 5	18	0.9	41	10	US-09-955-462A-1
C 6	17	0.8	30	11	US-09-439-429-18
C 7	17	0.8	34	11	US-09-907-907A-51
C 8	16	0.8	28	14	US-10-085-906-147
C 9	16	0.8	34	14	US-10-085-906-126
C 10	16	0.8	40	10	US-09-780-929-58
C 11	16	0.8	43	12	US-10-032-585-1704
C 12	15	0.7	17	12	US-10-238-700-1268
C 13	15	0.7	20	9	US-09-918-686-77
C 14	15	0.7	20	11	US-09-898-556A-85
C 15	15	0.7	20	12	US-10-353-150-77
C 16	15	0.7	25	14	US-10-098-263B-41465

17	15	0.7	25	14	US-10-098-263B-42093	Sequence 42093, A
18	15	0.7	25	14	US-10-098-263B-60871	Sequence 60871, A
c 19	15	0.7	25	14	US-10-098-263B-71610	Sequence 71610, A
20	15	0.7	25	14	US-10-098-263B-115355	Sequence 115355,
21	15	0.7	25	14	US-10-005-955-1249	Sequence 1249, Ap
c 22	15	0.7	27	10	US-09-263-959-785	Sequence 785, App
c 23	15	0.7	27	10	US-09-957-997-13	Sequence 13, Appl
c 24	15	0.7	32	12	US-10-091-281-317	Sequence 317, App
25	15	0.7	43	12	US-10-032-585-673	Sequence 673, App
26	15	0.7	43	12	US-10-032-585-1831	Sequence 1831, Ap
27	15	0.7	43	12	US-10-032-585-1868	Sequence 1868, Ap
28	15	0.7	43	11	US-09-993-346-18	Sequence 18, Appl
c 29	14	0.7	14	10	US-09-263-959-669	Sequence 669, App
30	14	0.7	15	10	US-09-263-959-622	Sequence 622, App
31	14	0.7	15	10	US-09-263-959-627	Sequence 627, App
32	14	0.7	15	10	US-09-263-959-650	Sequence 650, App
33	14	0.7	15	10	US-09-263-959-933	Sequence 933, App
34	14	0.7	15	12	US-09-842-347-13	Sequence 13, Appl
35	14	0.7	15	14	US-10-287-919-1255	Sequence 1255, Ap
36	14	0.7	15	14	US-10-287-919-1342	Sequence 1342, Ap
37	14	0.7	16	10	US-09-263-959-720	Sequence 720, App
c 38	14	0.7	16	10	US-09-263-959-786	Sequence 786, App
39	14	0.7	16	10	US-09-263-959-821	Sequence 821, App
40	14	0.7	16	11	US-09-232-785-361	Sequence 361, App
41	14	0.7	16	11	US-09-232-785-362	Sequence 362, App
42	14	0.7	16	13	US-10-001-835-18	Sequence 18, Appl
43	14	0.7	17	9	US-09-866-108-65	Sequence 65, Appl
44	14	0.7	17	9	US-09-866-108-66	Sequence 66, Appl
45	14	0.7	17	9	US-09-866-108-67	Sequence 67, Appl

ALIGNMENTS

RESULT 1
US-10-085-906-6/c
; Sequence 6, Application US/10085906
; Publication No. US20030054371A1
; GENERAL INFORMATION:
; APPLICANT: Ying, Vincent
; APPLICANT: Wu, Paul
; APPLICANT: Gray, Gary S.
; TITLE OF INVENTION: POLYMORPHIC ELEMENTS IN THE
; TITLE OF INVENTION: COSTIMULATORY RECEPTOR LOCUS AND USES THEREOF
; FILE REFERENCE: GNN-5343CP2
; CURRENT APPLICATION NUMBER: US/10/085,906
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 60/126,215
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 09/534,061
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: PCT/US00/07938
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 545
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-906-6

Query Match 1.0%; Score 20; DB 14; Length 30;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 378 AAAAATAAATAATAATAAA 397
|||||
Db 30 AAAAATAAATAATAATAAA 11

RESULT 2
US-10-224-005-6/c
; Sequence 6, Application US/10224005

```
; Publication No. US20030143732A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: McSwiggen, James
; APPLICANT: Fosnaugh, Kathy
; TITLE OF INVENTION: RNA Interference Mediated Inhibition of Adenosine A1 Receptor (A1)
; TITLE OF INVENTION: Gene Expression Using Short Interfering RNA
; FILE REFERENCE: 900/041 (WBHB01-1110-A)
; CURRENT APPLICATION NUMBER: US/10/224,005
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: US 60/315,315
; PRIOR FILING DATE: 2001-08-28
; NUMBER OF SEQ ID NOS: 347
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Target sequence/siRNA sense
US-10-224-005-6

Query Match          0.9%; Score 18; DB 12; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2023 CTCACGGGATTCCTTCCA 2040
DB 19 CTCACGGGATTCCTTCCA 2

RESULT 3
US-10-224-005-167
; Sequence 167, Application US/10224005
; Publication No. US20030143732A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: McSwiggen, James
; APPLICANT: Fosnaugh, Kathy
; TITLE OF INVENTION: RNA Interference Mediated Inhibition of Adenosine A1 Receptor (A1)
; TITLE OF INVENTION: Gene Expression Using Short Interfering RNA
; FILE REFERENCE: 900/041 (WBHB01-1110-A)
; CURRENT APPLICATION NUMBER: US/10/224,005
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: US 60/315,315
; PRIOR FILING DATE: 2001-08-28
; NUMBER OF SEQ ID NOS: 347
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 167
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: siNA antisense region
US-10-224-005-167

Query Match          0.9%; Score 18; DB 12; Length 19;
Best Local Similarity 72.2%; Pred. No. 2.2e+02;
Matches 13; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2023 CTCACGGGATTCCTTCCA 2040
DB 1 CUCCAGGGAUUCUCCA 18

RESULT 4
US-09-955-462A-2/c
; Sequence 2, Application US/09955462A
; Patent No. US20020150913A1
; GENERAL INFORMATION:
; APPLICANT: Wilusz, Jeffrey
; APPLICANT: Wilusz, Carol
; APPLICANT: Gao, Min
; TITLE OF INVENTION: ANTISENSE OLIGODEOXYNUCLEOTIDES REGULATING EXPRESSION
```

```
; TITLE OF INVENTION: Compositions and Methods for Reproducing and Modulating Mammalian
; TITLE OF INVENTION: Messenger RNA Decapping
; FILE REFERENCE: 601-1-109N
; CURRENT APPLICATION NUMBER: US/09/955,462A
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/233,682
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 34
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: described in specification page 29
US-09-955-462A-2

Query Match          0.9%; Score 18; DB 10; Length 34;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 AAATAAATAAATAATAAA 397
DB 33 AAATAAATAAATAATAAA 16

RESULT 5
US-09-955-462A-1/c
; Sequence 1, Application US/09955462A
; Patent No. US20020150913A1
; GENERAL INFORMATION:
; APPLICANT: Wilusz, Jeffrey
; APPLICANT: Wilusz, Carol
; APPLICANT: Gao, Min
; TITLE OF INVENTION: Compositions and Methods for Reproducing and Modulating Mammalian
; TITLE OF INVENTION: Messenger RNA Decapping
; FILE REFERENCE: 601-1-109N
; CURRENT APPLICATION NUMBER: US/09/955,462A
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/233,682
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: described in specification page 28
US-09-955-462A-1

Query Match          0.9%; Score 18; DB 10; Length 41;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 AAATAAATAAATAATAAA 397
DB 37 AAATAAATAAATAATAAA 20

RESULT 6
US-09-439-429-18
; Sequence 18, Application US/09439429
; Publication No. US20030083275A1
; GENERAL INFORMATION:
; APPLICANT: Power, Christopher
; APPLICANT: Mayne, Michael B.
; TITLE OF INVENTION: ANTISENSE OLIGODEOXYNUCLEOTIDES REGULATING EXPRESSION
; FILE REFERENCE: 3045.00002
; CURRENT APPLICATION NUMBER: US/09/439,429
; CURRENT FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 60/062,718
```


; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
US-09-780-929-58

Query Match 0.8%; Score 16; DB 10; Length 40;
Best Local Similarity 68.8%; Pred. No. 2.4e+03;
Matches 11; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1059 GCATTGTGAGAACTTCA 1074
|||||:|||||:|||||
DB 8 GCAUUUGAGAACUCCA 23

RESULT 11

US-10-032-585-1704/c
; Sequence 1704, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1704
; LENGTH: 43
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-032-585-1704

Query Match 0.8%; Score 16; DB 12; Length 43;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 AAAATAAATAATAA 393
|||||:|||||:|||||
DB 22 AAAATAAATAATAA 7

RESULT 12

US-10-238-700-1268/c
; Sequence 1268, Application US/10238700
; Publication No. US20030153521A1
; GENERAL INFORMATION:
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: Nucleic Acid Treatment of Diseases or Conditions Related to Level
; FILE REFERENCE: 400/057 (MBH501-1158-A)
; CURRENT APPLICATION NUMBER: US/10/238,700
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: PCT/US 02/16840
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: US 60/318,471
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 4666
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1268
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-238-700-1268

Query Match 0.7%; Score 15; DB 12; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.6e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 AAAATAAATAATAA 394
|||||:|||||:|||||
DB 16 AAAATAAATAATAA 2

RESULT 13

US-09-918-686-77
; Sequence 77, Application US/09918686
; Patent No. US20020076720A1
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary
; APPLICANT: Prohl, Sean
; APPLICANT: Paepker, Bryan
; APPLICANT: Staehling-Hampton, Karen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: GENOMIC DELETIONS
; FILE REFERENCE: 240083.515
; CURRENT APPLICATION NUMBER: US/09/918,686
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR primer
US-09-918-686-77

Query Match 0.7%; Score 15; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.6e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1237 TGGGATTACAGGTGT 1251
|||||:|||||:|||||
DB 2 TGGGATTACAGGTGT 16

RESULT 14

US-09-898-556A-85/c
; Sequence 85, Application US/09898556A
; Publication No. US20030087849A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freiler
; TITLE OF INVENTION: ANTISENSE MODULATION OF HKR1 EXPRESSION
; FILE REFERENCE: RTS-0248
; CURRENT APPLICATION NUMBER: US/09/898,556A
; CURRENT FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 85
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-898-556A-85

Query Match 0.7%; Score 15; DB 11; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.6e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1234 TGTGGGATTACAGG 1248
|||||:|||||:|||||
DB 18 TGTGGGATTACAGG 4

RESULT 15

US-10-353-150-77
; Sequence 77, Application US/10353150
; Publication No. US20030157543A1
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Prohl, Sean
; APPLICANT: Paepker, Bryan
; APPLICANT: Staehling-Hampton, Karen
; -TITLE OF INVENTION: METHODS FOR IDENTIFYING

; TITLE OF INVENTION: GENOMIC DELETIONS
; FILE REFERENCE: 240083.515C1
; CURRENT APPLICATION NUMBER: US/10/353,150
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR primer
US-10-353-150-77

Query Match 0.7%; Score 15; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.6e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1237 TGGGATTACAGGTGT 1251
||| ||||| ||||| |||||
Db 2 TGGGATTACAGGTGT 16

Search completed: November 14, 2003, 19:28:16
Job time : 628.131 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 03:50:47 ; Search time 8568.5 Seconds
(without alignments)
11258.085 Million cell updates/sec

Title: US-10-005-337A-1

Perfect score: 2358

Sequence: 1 ggaatccttcattgtttaaca.....caggtcgaggccaccatgg 2358

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rnd.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2358	100.0	2358	6	AX468603	AX468603 Sequence
2	2301.8	97.6	142902	2	AC119234	AC119234 Mus muscu
3	879.8	37.3	229640	2	AC105469	AC105469 Rattus no
4	879.8	37.3	238344	2	AC097115	AC097115 Rattus no
C 5	804.8	34.1	838	11	BV062893	BV062893 S212P6025
6	646.2	27.4	723	10	AF478692	AF478692 Mus muscu
7	640.2	27.2	665	11	BV076484	BV076484 S212P6036
8	367.8	15.6	2074	6	AX468604	AX468604 Sequence
9	367.8	15.6	2074	9	AF131884	AF131884 Homo sapi
C 10	347.6	14.7	50111	9	AL590622	AL590622 Human DNA
11	128.6	5.5	1901	6	AX322775	AX322775 Sequence
12	128.6	5.5	1901	6	BD094076	BD094076 Shear str
13	128.6	5.5	1901	9	HSRNCACINP	X83703 H.sapiens m
14	128.6	5.5	1901	11	G28603	G28603 human STS S
C 15	127.2	5.4	160350	2	AC074094	AC074094 Homo sapi
16	125.8	5.3	110480	10	AC122467	AC122467 Mus muscu
17	125.8	5.3	169126	2	AC132348	AC132348 Mus muscu
C 18	125.2	5.3	152346	2	AC102022	AC102022 Mus muscu
C 19	125.2	5.3	258445	2	AC122205	AC122205 Mus muscu
C 20	122.4	5.2	158357	9	AL365434	AL365434 Human DNA
C 21	121.6	5.2	1940	4	AF131883	AF131883 Oryctolag
C 22	119.6	5.1	25203	10	AL365324	AL365324 Mouse DNA
C 23	119.6	5.1	184754	2	AC022675	AC022675 Mus muscu
C 24	119	5.0	234469	2	AC119697	AC119697 Rattus no
C 25	117.8	5.0	153899	10	AL772285	AL772285 Mouse DNA
C 26	117.8	5.0	191485	2	AC091467	AC091467 Mus muscu
C 27	117.8	5.0	227165	10	AL626768	AL626768 Mouse DNA
C 28	117.8	5.0	243290	10	AL663088	AL663088 Mouse DNA
C 29	117.8	5.0	317209	2	AL772175	AL772175 Mus muscu
C 30	116.6	4.9	167591	2	AC136093	AC136093 Rattus no
C 31	116.6	4.9	179706	2	AC128045	AC128045 Rattus no
C 32	116.6	4.9	231573	2	AC123610	AC123610 Mus muscu
C 33	116.6	4.9	245489	2	AC105665	AC105665 Rattus no
C 34	116.6	4.9	267971	2	AC128995	AC128995 Rattus no
C 35	116.4	4.9	183205	2	AC133734	AC133734 Rattus no
C 36	116.4	4.9	238313	2	AC109696	AC109696 Rattus no
C 37	116.4	4.9	240006	2	AC094069	AC094069 Rattus no
C 38	116.2	4.9	249982	2	AC108661	AC108661 Rattus no
C 39	115.8	4.9	167223	2	AC113514	AC113514 Mus muscu
C 40	115.8	4.9	219619	2	AC116128	AC116128 Mus muscu
C 41	115.4	4.9	244328	2	AC094938	AC094938 Rattus no
C 42	115.2	4.9	171403	10	AC124423	AC124423 Mus muscu
C 43	114.8	4.9	294703	2	AC103396	AC103396 Mus muscu
C 44	114.6	4.9	176926	2	AC138341	AC138341 Mus muscu
C 45	114.4	4.9	167601	2	AC115886	AC115886 Mus muscu

ALIGNMENTS

RESULT 1
AX468603

LOCUS

DEFINITION

AX468603

SEQUENCE

AX468603.1

VERSION

AX468603.1

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1

Schwartz, B., Branellec, D. and Chien, K.

Sequences upstream of the carp gene, vectors containing them and

uses thereof

AX468603 Sequence 1 from Patent WO0246220. 2358 bp DNA linear PAT 16-JUL-2002

JOURNAL		Patent: WO 0246220-A 1 13-JUN-2002;			
FEATURES		Aventis Pharma S.A. (FR) ; The Regents of The University of California at San Diego (US) ; Benoit, Patrick (FR)			
source	Location/Qualifiers				
	1. 2358				
	/organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090"				
BASE COUNT		636 a	567 c	542 g	613 t
ORIGIN					
Query Match		100.0%;	Score 2358;	DB 6;	Length 2358;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2358;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GGATCTTTTCATGTTTAA	CAATATCA	CCCTAAC	CCGAGGGAACGCGCTGCACAGT 60
Db	1	GGATCTTTTCATGTTTAA	CAATATCA	CCCTAAC	CCGAGGGAACGCGCTGCACAGT 60
Qy	61	GGCTTTGCCACCCATGAT	CTTCTAC	AGGTGCT	CAATGGGAATTTCTGGAGCTTCTCT 180
Db	61	GGCTTTGCCACCCATGAT	CTTCTAC	AGGTGCT	CAATGGGAATTTCTGGAGCTTCTCT 180
Qy	121	ACACTTCTGCAAGCCCAT	CTCTACA	AGGTGCT	CAATGGGAATTTCTGGAGCTTCTCT 180
Db	121	ACACTTCTGCAAGCCCAT	CTCTACA	AGGTGCT	CAATGGGAATTTCTGGAGCTTCTCT 180
Qy	181	TTTCAGATCAGGCTGAT	TTCTAG	GGCAGCAGT	TTCTCAACCTGGGGGCTCGACCCCTTTGG 240
Db	181	TTTCAGATCAGGCTGAT	TTCTAG	GGCAGCAGT	TTCTCAACCTGGGGGCTCGACCCCTTTGG 240
Qy	241	GGGATCAAAACGACCTT	TACAGG	GGTCA	CATATCTATATGTCAGGATTTTA 300
Db	241	GGGATCAAAACGACCTT	TACAGG	GGTCA	CATATCTATATGTCAGGATTTTA 300
Qy	301	CATTACGATTCGTAA	CGTAG	CAAAATTA	CAGGTATGAATAGCAATGAATATTTAT 360
Db	301	CATTACGATTCGTAA	CGTAG	CAAAATTA	CAGGTATGAATAGCAATGAATATTTAT 360
Qy	361	GATTGAAGTCAACCA	CAATAG	AGCGCCG	CACTGTTCTAGAGAAATTCACCTGGGTG 420
Db	361	GATTGAAGTCAACCA	CAATAG	AGCGCCG	CACTGTTCTAGAGAAATTCACCTGGGTG 420
Qy	421	GGGAAGGTTTGGAA	AGGCTTT	CTGTC	CAATCTTCAATCTTCAAAAGTATGTTTCA 480
Db	421	GGGAAGGTTTGGAA	AGGCTTT	CTGTC	CAATCTTCAATCTTCAAAAGTATGTTTCA 480
Qy	481	GAAAGCCTTTCAG	CTGTTCT	GCTGGG	CTCTAGTAAGTCTGAGTAGGAACTGTATGTAC 540
Db	481	GAAAGCCTTTCAG	CTGTTCT	GCTGGG	CTCTAGTAAGTCTGAGTAGGAACTGTATGTAC 540
Qy	541	CAGGCTCGTCTTAT	TGGTGG	AGCCAA	GCATCGTGGGTGGAGCGAAGCGCAACCT 600
Db	541	CAGGCTCGTCTTAT	TGGTGG	AGCCAA	GCATCGTGGGTGGAGCGAAGCGCAACCT 600
Qy	601	CACCTTCTAGCT	CTGCAT	AGCAAG	TAGCTAATGTTTCTGCTCTAGGTGTCATCT 660
Db	601	CACCTTCTAGCT	CTGCAT	AGCAAG	TAGCTAATGTTTCTGCTCTAGGTGTCATCT 660
Qy	661	CTGTGAATCGAG	ATCCTT	GGCCCTT	GTGAATTTAGGGAGGCACAAAATACTCAGAGATTTC 720
Db	661	CTGTGAATCGAG	ATCCTT	GGCCCTT	GTGAATTTAGGGAGGCACAAAATACTCAGAGATTTC 720
Qy	721	AAGACTGCTCAG	CGCCAG	AGTCTCT	CTTCTCAAGGAAAGGTCTCAACTCTCAGCCCCC 780
Db	721	AAGACTGCTCAG	CGCCAG	AGTCTCT	CTTCTCAAGGAAAGGTCTCAACTCTCAGCCCCC 780
Qy	781	TTAGCTCTGAG	CTCAG	GCCTT	GGAAACAAACCGGCCACAGGAATGAGAAAGCTGCCATAGCTG 840
Db	781	TTAGCTCTGAG	CTCAG	GCCTT	GGAAACAAACCGGCCACAGGAATGAGAAAGCTGCCATAGCTG 840
Qy	841	CTTGTCACTTCA	AGAGGT	CAAGAAA	ATAGTGTAAACCATGAAACGAGAACCAACAG 900
Db	841	CTTGTCACTTCA	AGAGGT	CAAGAAA	ATAGTGTAAACCATGAAACGAGAACCAACAG 900

Db	841	CTTGTCACTTCA	AGAGGT	CAAGAAA	ATAGTGTAAACCATGAAACGAGAACCAACAG 900
Qy	901	TTATCCATTGTAG	CGTCTC	TAGGAC	GATAGGACAGAGAACACTAGGAGGGGAACC 960
Db	901	TTATCCATTGTAG	CGTCTC	TAGGAC	GATAGGACAGAGAACACTAGGAGGGGAACC 960
Qy	961	CACGAAGGAC	AGGTAT	TAGTGT	TTTTCAGGCAATGCTTTGTACTGAAGATTCT 1020
Db	961	CACGAAGGAC	AGGTAT	TAGTGT	TTTTCAGGCAATGCTTTGTACTGAAGATTCT 1020
Qy	1021	AGAAACA	CAATTT	TGCTG	TGAACAGCTGAAGTGGGGTGGTCTTACCCCATGTTCA 1080
Db	1021	AGAAACA	CAATTT	TGCTG	TGAACAGCTGAAGTGGGGTGGTCTTACCCCATGTTCA 1080
Qy	1081	TGGAGGGGT	GAGTGAGG	GAGACAT	ATATATGATGCCAGCATATAACAACATACACACA 1140
Db	1081	TGGAGGGGT	GAGTGAGG	GAGACAT	ATATATGATGCCAGCATATAACAACATACACACA 1140
Qy	1141	CCCTAATTA	ACAATTT	CCCTCT	TCTACTGACACCCCTTCACTCTCTCTTTTCAAAAAA 1200
Db	1141	CCCTAATTA	ACAATTT	CCCTCT	TCTACTGACACCCCTTCACTCTCTCTTTTCAAAAAA 1200
Qy	1201	TAAAAA	AGTAT	TTTAT	TGCTCTTA CGATAGAATCTTTTCTCGAACTATAAAAAGATC 1260
Db	1201	TAAAAA	AGTAT	TTTAT	TGCTCTTA CGATAGAATCTTTTCTCGAACTATAAAAAGATC 1260
Qy	1261	TAAATAT	TATATTT	TTCACAT	TTTAAATCTTTAGTGGGCTCTTTTGTTCGGTGTAGGAATAGAACA 1320
Db	1261	TAAATAT	TATATTT	TTCACAT	TTTAAATCTTTAGTGGGCTCTTTTGTTCGGTGTAGGAATAGAACA 1320
Qy	1321	TTTTGCC	CTCTCA	ACAGCA	AAAGCTTTGGGGCTCTTTTGTTCGGTGTAGGAATAGAACA 1380
Db	1321	TTTTGCC	CTCTCA	ACAGCA	AAAGCTTTGGGGCTCTTTTGTTCGGTGTAGGAATAGAACA 1380
Qy	1381	CGAGAG	CCCCG	TGATCT	ATAGGAGATCTATATATAGGCCATAGTCTCCAGCCTCAG 1440
Db	1381	CGAGAG	CCCCG	TGATCT	ATAGGAGATCTATATATAGGCCATAGTCTCCAGCCTCAG 1440
Qy	1441	AGGCA	CAATTTT	CTCGG	CTCTTAAAGCTTTTCCACAGCATTTGGAACCTTTTACTGAC 1500
Db	1441	AGGCA	CAATTTT	CTCGG	CTCTTAAAGCTTTTCCACAGCATTTGGAACCTTTTACTGAC 1500
Qy	1501	AGCAT	CCAA	GTGTCT	CTGCTAAGAACTGGAGCTCACTCTCTCTGTGCATCCTCCG 1560
Db	1501	AGCAT	CCAA	GTGTCT	CTGCTAAGAACTGGAGCTCACTCTCTCTGTGCATCCTCCG 1560
Qy	1561	CCGCT	TTTGGG	TAGAT	CTCTGATTTAGAAACAGCGTGAGCCTGGT 1620
Db	1561	CCGCT	TTTGGG	TAGAT	CTCTGATTTAGAAACAGCGTGAGCCTGGT 1620
Qy	1621	GCAC	TAA	TTAT	TGGCCAGTGACACCATAGAGTCAAAAGTCATTTACTGAAATGCTTTCAATT 1680
Db	1621	GCAC	TAA	TTAT	TGGCCAGTGACACCATAGAGTCAAAAGTCATTTACTGAAATGCTTTCAATT 1680
Qy	1681	CTCT	TAAT	AGCT	CGATGGCATGTCA CAGGGCCATTTTAGCTGCAGACATCCTCCAG 1740
Db	1681	CTCT	TAAT	AGCT	CGATGGCATGTCA CAGGGCCATTTTAGCTGCAGACATCCTCCAG 1740
Qy	1741	AGAA	TTTCC	AAA	CAGATAGACAAGTGGCA CCGACCCCATCTCTCTTCCCTCGGGCTGA 1800
Db	1741	AGAA	TTTCC	AAA	CAGATAGACAAGTGGCA CCGACCCCATCTCTCTTCCCTCGGGCTGA 1800
Qy	1801	TTAT	CCCC	GAA	ATAGGATGTCCCAAGCAACACTTCCAGCCCACTGGAGTCTCTATAA 1860
Db	1801	TTAT	CCCC	GAA	ATAGGATGTCCCAAGCAACACTTCCAGCCCACTGGAGTCTCTATAA 1860
Qy	1861	GTCC	AGT	TTAT	CAGAAAGATATGGCTGTAAAGTGTGATGCAAGTCTTGTGATTTTCTTGAT 1920
Db	1861	GTCC	AGT	TTAT	CAGAAAGATATGGCTGTAAAGTGTGATGCAAGTCTTGTGATTTTCTTGAT 1920
Qy	1921	ACGT	TAG	TCAT	ATGAGAGCTGACAAAGAGGAAAAAGACGAGATGTGGTGCATATTTA 1980
Db	1921	ACGT	TAG	TCAT	ATGAGAGCTGACAAAGAGGAAAAAGACGAGATGTGGTGCATATTTA 1980


```
/note="clone boundary
clone_end:T7
site:
end_sequence:RWXP57TJB"
complement(7252..8118)
/note="clone boundary
clone_end:T7
site:
end_sequence:RWXP57TJC"
213542..215467
/note="wgs_end_extension
clone_end:T7"
215568..216791
/note="wgs_end_extension
clone_end:T7"
BASE COUNT 61330 a 47195 c 48097 g 63316 t 9702 others
ORIGIN
Query Match 37.3%; Score 879.8; DB 2; Length 229640;
Best Local Similarity 82.3%; Pred. No. 1e-245;
Matches 1209; Conservative 0; Mismatches 177; Indels 83; Gaps 14;
QY 879 ATGAAACGAGAGAACCAACAGTATCCATTGATACGCTCTCAGACAGATAGGACAGAG 938
DB 140789 ATGAAACGAGAGAACCAACAGTATCCATTGATACGCTCTCAGACAGATAGGACAGAG 140848
QY 939 AGAACACTAGGAGGGGAAACCAACGAGCAAGGTATAGTGTGTTTTCAGGCG 998
DB 140849 ----CACTAGGAGAGAGAACCAACGAGCAAGGTATAGTGTGTTTTCAGGCG 140899
QY 999 AATGCTCTGCTGAGTATCTAGAAACACAAATTCGCTGTTGAACAGCTGAAGTGGGT 1058
DB 140900 AATGCTCTGCTGAGTATCTAGAAACACAAATTCGCTGTTGAACAGCTGAAGTGGGT 140959
QY 1059 GGGG---GTTCCTACCCATGTTCTAGGAGGCTGAGTGAGGAGACAGATATATGATG 1115
DB 140960 GGGGTAAGAAATAACACACAGTCTACTGAAGGCTGAGGAGGAGACAGATATGATG 141019
QY 1116 GCGAGNTAACAAACATACACACACCTTAATTAACATTCCTCTTCTACTGACACCCC 1175
DB 141020 GCGAGCAT-----ACAAACATAACACCTTAATTAATGCTTCCTCTGCTACTGACACTCT 141074
QY 1176 CTTCACTCTCTCTTTCATA-----AAAAATAAAAAAAGT 1210
DB 141075 CCGTTCACTCTCTTTCATAAAAAATAAAAAAATAAATAAATAAATAAATAAATAAATAA 141134
QY 1211 ATTTATGTTGGCTTTACGATAGAAATCTTTCCTCGAATATATAAAGATCTAAATATTTA 1270
DB 141135 ATTTATGTTGGCTTTACGATAGAAATCTTTCCTCGAATATATAAAGATCTAAATATTTA 141189
QY 1271 TATTTTTCATTTTAAATATCTTAGCGATGACAGCCAGAAACAAAGTATTTTTCCTCT 1330
DB 141190 TATTTTTCATTTTAAATATCTTAGCGATGACAGCCAGAAACAAAGTATTTTTCCTCT 141248
QY 1331 CTCACAGCAAAAGCTTGGGCGCTT-TTGTGTTTCGTTTGGAGTAAGAACACAGAGAGCC 1389
DB 141249 CTCACAGCAAAAGCTTGGGCGCTT-TTGTGTTTCGTTTGGAGTAAGAACACAGAGAGCC 141299
QY 1390 CGTGTATCTAGGAGATGCTCTATCATATTAGCCCATGAGTCTCCAGCCTCAGAGCGCAT 1449
DB 141300 CATGTATCTAGGAGATGCTCTATCATATTAGCCCATGAGTCTCCAGCCTCAGAGCGCAT 141347
QY 1450 TTTCTCGGCTCTTAAAGCTTT-TCCACAGCAATTTGGAAACATTTTACTGACAGCATCA 1508
DB 141348 TTTCTCGGCTCTTAAAGCTTT-TCCACAGCAATTTGGAAACATTTTACTGACAGCATCA 141407
QY 1509 AGTTGTTCTCTGCTAGAACTGAGTCTCATCTCTCTGTCATCTCTGTCGCTCTGCGCGCTTT 1568
DB 141408 AGTTGTTCTCTGCTAGAACTGAGTCTCATCTCTCTGTCATCTCTGTCGCTCTGCGCGCTTT 141467
QY 1569 GGGGTAGATCTCTGATAGCTTTAGCTTTAGAAACACAGGCTGAGCCTGTGGTGCATAAT 1628
DB 141468 GAGGTCAATATCCGATTTAGCTTTCTGTTTATAGACACA-----TGCCATGCACTAAT 141520
```

```
1629 TATGCCAGTGACACCATAGAGTCAAGTGCATTACTGAATGCTTTCAATTTCTCTAAT 1688
DB 141521 TATGCCAGTGACACCATAGAGTCAAGTGCATTACTGAATGCTTTCAATTTCTCTAAT 141580
QY 1689 GCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1748
DB 141581 GATGTTAAGTGGCTGTCATGGGGCCATTTAGCCCTGGACATCACTCCAGAGAATCT 141640
QY 1749 AAACAGATAGAGACAAGTGGCCACCCAGACCCATCTCTCTCCCTCGGCTGATTTATCC 1808
DB 141641 GAACAGATATAGACAAGTGTACCAAGACCCCA-----CTTCCCTCCGCTGTTTATCC 141696
QY 1809 AGAATATAGATGCTCCCAAGCAACACTTCCAGACCAACTGGAGTGTCTGATAAGTCCAG 1868
DB 141697 AGAATATAGATGCTCCCAAGCAACACTTCCAGACCAACTGGAGTGTCTGATAAGTCCAG 141756
QY 1869 ATCAGAAAGATATGCTGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1928
DB 141757 ATCAGAAAGATATGCTGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 141816
QY 1929 CATATGAGAGCTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1988
DB 141817 CATACGAAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 141876
QY 1989 CTGTCCTCTGCTTCCCGATAGCTGCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2048
DB 141877 CTGT-CCCTGCTTCCCGATAGCTGCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 141935
QY 2049 CAAAGGAATGACCTCTCACATTTCTCTGATTCGCATACGCGCGGCGGCGGCGGCGGCGG 2108
DB 141936 CAATGGAATGCGCTCTCACATTTCTCTGATTCGCATACGCGGCGGCGGCGGCGGCGG 141995
QY 2109 CTCCCTCTTGGCTTCCAGACACACTAAGTCTGGAATGAAATTCACCTGCTCTGAAATG 2168
DB 141996 CTCCCTCTTGGCTTCCAGACACACTAAGTCTGGAATGAAATTCACCTGCTCTGAAATG 142055
QY 2169 GCGACTGTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2228
DB 142056 GCGACAGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 142115
QY 2229 CCTAGCTATATACGCGGCTGCTGAGGCGGCTCCACAGGCGGCGGCTCCAGGCGGCTCAT 2288
DB 142116 TCCTAGCTATATACGCGGCTGCTGAGGCGGCTCCACAGGCGGCGGCTCCAGGCGGCTCAG 142175
QY 2289 CCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2317
DB 142176 CCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 142204
```

RESULT 4
AC097115
LOCUS
DEFINITION Rattus norvegicus clone CH230-26A2, *** SEQUENCING IN PROGRESS ***,
2 unordered pieces.
AC097115
VERSION AC097115.6 GI:24956605
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 238344)
Muzny, D., Metzker, M., Lee, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,
Balwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P., Fraser, C.M., Gabisi, A., Ganca, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerrero, W., Gunaratne, P., Haaland, W., Hamilton, C., Hamilton, C., Hamiton, K., Harvey, X., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hollins, B., Howells, S., Hui, J., Hodgson, A., Hogues, M., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnston, R., Jollivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.B., Lebowitz, H., Levay, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.J., Lorensushewa, L., Louisedged, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norrie, S., Nwaokelimeh, O., Okwuonu, G., Olarnpunagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Fuazo, M., Quintero, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shateman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, J., Steinkamp, M., Strong, R., Sutton, A., Svatek, A., Taborski, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wlaczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

Direct Submission
Unpublished

2 (bases 1 to 238344)
Worley, K.C.
Submitted (11-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 238344)

Rat Genome Sequencing Consortium.
Rat Submissions
Submitted (14-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 14, 2002 this sequence version replaced gi:22855482.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly ('contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genomex Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----

```
QY 1390 CBTGTATCTAGGCAGATGCTCTATCATATTAGCATGAGTCTCAGGCTCAGAGCACATTT 1449
Db 26929 CATGTATCTAGGCAGAT-----GCTATGAATCTCAGGCTCAAGTGCATGTT 26976
QY 1450 TTTCTCGGGCTCTTTAAGCTTT-TCCACAGCAATTTGGAAACTTTTACTGACAGCATCA 1508
Db 26977 TTTCTCAGGCTCTCCAGGCTTTCCGCCACAGCAATTTGAAAGCTTTTACTGACAGGTC 27036
QY 1509 AGTTGTGCTTCTGCTAAGAACTGGAACTCATCTCTCTGTGCATCACTTCGGCCCGCTTT 1568
Db 27037 AGTCGTGCTTCTGCTAATAACTGGGCTCATATGTCTCTGACCACTCTGACCCCATTT 27096
QY 1569 GGGGTAGATCCTCTGATTAAGCTTTCAGATTTAGAACACGGTGAGCCTGTGTGCATTA 1628
Db 27097 GAGGTCAATATTCGGATTAAGCTTCTGTTTTAGACACA-----TGCCATGCACTAAT 27149
QY 1629 TATGCCAGTGACACCATAGAGTCAAGTGCATTACTGAATGCTTTCAATTTCTCCTAAT 1688
Db 27150 TATGCCAGTGACACCATAAAGTAAAGTGCACTACTGAATGCTTTCAATTTCTCATAAT 27209
QY 1689 GCTGTACATGATGCTGTACAGGGCCATTTTGTGTCGACATCACTCCAGAGATTCCT 1748
Db 27210 GATGTAAAGTGTGCGTGTCTATGGGGCAATTTAGCCCTGGACATCACTCCAGAGATTC 27269
QY 1749 AAACAGATAGACAAAGTGGCACCAGCCCATCTCTCTCCCTCGGGCTGATTCCTCC 1808
Db 27270 GAACAGATATAGACAAAGTGTACCAAGACCCA-----CTTCCCCCGGGCTGTTATTC 27325
QY 1809 AGAAATAGGATGTCCCAAGCAACACTTCCAGCCAACTCGAGTGTGTATAGTCAGATT 1868
Db 27326 AGGAATAGATGTCCCAAGCAACACTTCCAGCCAACTCGAGTGTGTATAGCCAGATT 27385
QY 1869 ATCAGAAAGATATGGCTGTAAAGTGTGATGCACAGTCTTGCAATTTCTTGTATACGTT 1928
Db 27386 ATCAGAAAGATATTTGCGTGAAGTGTGATGCACAACTCTTGCACTTTCTGATAGGTT 27445
QY 1929 CATATGAGAGTGTACAAAGAGGAAAGAGGAGCGAGTGTGTCATATTAACAGGCAG 1988
Db 27446 CATACGAAAGTGTACAGAGAGGAGGAGGAGCGAGTGTGTCATATTAACAGGCAG 27505
QY 1989 CTGTCCCTCTGGGTTCCCGATACGTGGGATGACTCGCATTTGCTGAGCGGTGTGCTACT 2048
Db 27506 CTGT-CCCTGGCTTCCCGATAGTAGATGACTCGCATTTGCTGGGCGGTGTGCTACTGC 27564
QY 2049 CAAAGGAATGACCTCTCAATTTCTCTGATTCGCATFACGCGCGGCGGCTTGTCTAT 2108
Db 27565 CAATGGAATGGCCCTCTCAATTTCTCTGATTCGCATFACGCGCGGCGGCTTGTCTAT 27624
QY 2109 CTCCCTCTTGGGCTTCCAGACACTAAGTCTGGAATGAAATTCACCTGCCTCTGAATTG 2168
Db 27625 CTCCCTCTTGGGCTTCCAGACACTAAGTCTGGAATGAAATTCACCTGCCTCTGAATTG 27684
QY 2169 GCACCTGGTGGGCGGAGGGGTGTGACTTGGCTTCCAGGCTGGAAGATTAATCTACCC 2228
Db 27685 GCACAGCGGAGGCGAGGGGTGTGACTTGGCTTCCAGGCTGGAAGATTAATCTACCC 27744
QY 2229 CCTAGCTATATAACGGGCTGTGTGGAGGGGCTCCACAGGGCGCAATTCAGGGGTTCAT 2288
Db 27745 TCCTAGCTATATAAGAGGCTGGGCTGGAGGGGCTCCACAGGGCGCAATTCAGGGGTTCAG 27804
QY 2289 CCACAGAGAGAGAAACATAGACTCGAGG 2317
Db 27805 CCACAGAGAGAGAGAAACATAGACTCACGG 27833
```

```
RESULT 5
LOCUS BV062893/C
DEFINITION S212P05252PC3.T0 CZECHII/E1 Mus musculus STS genomic, sequence
tagged site.
ACCESSION BV062893
VERSION BV062893.1 GI:31178688
```

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 838)

AUTHORS Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C., Lander,E.S., Lindblad-Toh,K. and Daly,M.J.

TITLE The mosaic structure of variation in the laboratory mouse genome

JOURNAL Nature 420 (6915), 574-578 (2002)

MEDLINE 22354684

PUBMED 12466852

COMMENT

Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 838

Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS reads were placed uniquely on the MGSv3, C57BL/6J assembly and SNP detection was carried out by SSAHA-SNP. 225,000 reads were annotated

as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J and the strain from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.

FEATURES

source

1..838

Location/Qualifiers

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="CZECHII/E1"

/db_xref="taxon:10090"

/map="19 27-787 35750684-35751444"

/clone_lib="CZECHII/E1"

<1..838

STS

BASE COUNT 235 a 164 c 208 g 231 t

ORIGIN

Query Match 34.1%; Score 804.8; DB 11; Length 838;
Best Local Similarity 98.9%; Pred. No. 5.2e-224;
Matches 831; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY 1004 CTTGTACTGAAGATTTAGAAAACAAATTTGCTGTTGAACAGCTGAAGTGGGGTGGGG 1063

Db 838 CTCGTTCTGAAGGTTCTAGAAAACAAATTTGCTGG-TGAACAGCTGAAGTGGGGTGGGG 780

QY 1064 TTCTTACCCCATGTTTCATGGAAGGGTGAAGAGAGACAGATATATGATGCCAGCAT 1123

Db 779 TTCTTACCCCATGTTTCATGGAAGGGTGAAGAGAGACAGATATATGATGCCAGCAT 720

QY 1124 AACAAACATACACACACCTTAATTAACACTTCCCTCTTCTACTGCACACCCCTTCACTC 1183

Db 719 AACAAACATACACACACCTTAATTAACACTTCCCTCTTCTACTGCACACCCCTTCACTC 660

QY 1184 TCCTCTTTCATAAAAATAAAAAAGTATTTTATGTGGCTCTTACCATAGAAATCTTTCTCT 1243

Db 659 TCCTCTTTCATAAAAATAAAAAAGTATTTTATGTGGCTCTTACCATAGAAATCTTTCTCT 600

QY 1244 CGAACTATAAAAAGATCTAAATATTTATATTTTCACTTTTAATATCTTAGCGATGACA 1303

Db 599 CGAACTATAAAAAGATCTAAATATTTATATTTTCACTTTTAATATCTTAGCGATGACA 540

QY 1304 AGCCAGAAACAGATATTTTTCCTCTCTCACACAGAGCTTGGGGCTTTTGTTCCT 1363

Db 539 AGCCAGAAACAGATATTTTTCCTCTCTCACACAGAGCTTGGGGCTTTTGTTCCT 481

QY 1364 GTGTTAGGAATAGAACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1423


```
Db 480 GTGTTAGGAATAGAACACGAGAGCCCGGTGATCTTAGGCAGATGCTCTATCATTAGGCCA 421
Qy 1424 TGAGTCTCCAGCCTCAGACGACATATTTTCTCGGGCTCTCTTAAGCTTTTCCACAGCAT 1483
Db 420 TGAGTCTCCAGCCTCAGACGACATATTTTCTCGGGCTCTCTTAAGCTTTTCCACAGCAT 361
Qy 1484 TGGGAAACTTTACTGACAGCATCCAAAGTTGGCTTCTGCTAAGAACTGGACTCACATCTC 1543
Db 360 TGGGAAACTTTACTGACAGCATCCAAAGTTGGCTTCTGCTAAGAACTGGACTCACATCTC 301
Qy 1544 TCTGTGATCACTTCGGCCCGTTTTGGGTAGATCCTCTGATTAGCCTTCAGATTTAGAA 1603
Db 300 TCTGTGATCACTTCGGCCCGTTTTGGGTAGATCCTCTGATTAGCCTTCAGATTTAGAA 241
Qy 1604 CACGGTGAGCCTGTGGTGCACTAATATATGGCCAGTGACACCATAGAGTCAAAGTGCATTA 1663
Db 240 CACGGTGAGCCTGTGGTGCACTAATATATGGCCAGTGACACCATAGAGTCAAAGTGCATTA 181
Qy 1664 CTGAATGCTTCAATTTCTCTTAATGTGGTACGATGGCATGTCCAGGGCCATTTTACG 1723
Db 180 CTGAATGCTTCAATTTCTCTTAATGTGGTACGATGGCATGTCCAGGGCCATTTTACG 121
Qy 1724 TGCAGACATCACTCCAGAGATTCACAAACAGATAGACAGTGGCACCAGCCCATCT 1783
Db 120 TGCAGACATCACTCCAGAGATTCACAAACAGATAGACAGTGGCACCAGCCCATCT 61
Qy 1784 CTTTCCCTCGGGCTGATTTATCCCCAGAAATAGATGTCCCAAAGCAACACTTCCAGGCC 1843
Db 60 CTTTCCCTCGGGCTGATTTATCCCCAGAAATAGATGTCCCAAAGCAACACTTCCAGGCC 1

RESULT 6
AF478692 723 bp DNA linear ROD 10-OCT-2002
LOCUS Mus musculus cardiac ankyrin repeat protein (Carp) gene, promoter
DEFINITION
AF478692
ACCESSION
VERSION AF478692.1 GI:19110906
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Maeda,T., Sepulveda,J., Chen,H.H. and Stewart,A.F.R.
alpha1-Adrenergic activation of the cardiac ankyrin repeat protein
gene in cardiac myocytes
JOURNAL Gene 297 (1-2), 1-9 (2002)
REFERENCE 2 (bases 1 to 723)
AUTHORS Maeda,T., Sepulveda,J. and Stewart,A.F.R.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-2002) Cardiovascular Institute, University of
Pittsburgh, 200 Lothrop Street, Pittsburgh, PA 15213, USA
FEATURES
source
1. .723
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="FVB"
/db_xref="taxon:10090"
1. .>723
/gene="Carp"
1. .720
/note="contains promoter and 5' UTR"
<721. .>723
/gene="Carp"
/product="cardiac ankyrin repeat protein"
721. .>723
/gene="Carp"
/codon_start=1
/product="cardiac ankyrin repeat protein"
/protein_id="AAL85342.1"
```

```
/db_xref="GI:19110907"
BASE COUNT 187 a 185 c 180 g 171 t
ORIGIN
Query Match 27.4%; Score 646.2; DB 10; Length 723;
Best Local Similarity 97.5%; Pred. No. 1.5e-177;
Matches 699; Conservative 0; Mismatches 13; Indels 5; Gaps 4;
Qy 1601 GAACACGGTGAGCCTGTGGTGCACTAATATATGGCCAGTGACACCATAGAGTCAAAGTGC 1660
Db 1 GAACACGGTGAGCCTGTGGTGCACTAATATATGGCCAGTGACACCATAGAGTCAAAGTGC 60
Qy 1661 TTAAGTGAATGCTTTCAATTTCTCTTAATGTGGTACGATGGCATGTCCAGGGCCATTTT 1720
Db 61 TTAAGTGAATGCTTTCAATTTCTCTTAATGTGGTACGATGGCATGTCCAGGGCCATTTT 120
Qy 1721 AGTGCAGACATCACTCCAGAGATTCCAAACAGATAGACAAAGTGGCACCAGCCCA 1780
Db 121 AGTGCAGACATCACTCCAGAGATTCCAAACAGATAGACAAAGTGGCACCAGCCCA 180
Qy 1781 TCTCCTTCCCTCGGGCTGATTTATCCCCAGAAATAGATGTCCCAAAGCAACACTTCCCA 1840
Db 181 TCTCCTTCCCTCGGGCTGATTTATCCCCAGAAATA-GATGTCCCAAAGCA--ACTTCCCA 237
Qy 1841 GCCAACTGGAGTGTGATAAGTCCAGTTATCAGAAAGATATGGCTGTAAGTGTGATGCAC 1900
Db 238 GCCAACTGGAGTGTGATAAGTCCAGTTATCAGAAAGATATGGCTGTAAGTGTGATGCAC 297
Qy 1901 AGTGTGTGCAATTTCTTGATACGTTAGTATATAGAGAGCTGACAAAGAGGAAAAAGAGC 1960
Db 298 AGTGTGTGCAATTTCTTGATACGTTAGTATATAGAGAGCTGACAAAGAGG-AAAAGAGC 356
Qy 1961 AGCGATGTGTGCAATATTAACAGGAGCTGTCCCTGGCTTCCCGATAGTGGGATGCAC 2020
Db 357 AGCGATGT-GTGCAATATTAACAGGAGCTGTCCCTGGCTTCCCGATAGTGGGATGCAC 415
Qy 2021 TCGCATTTGCTGAGCGGTGTGGTCACTGCCAAAGGAATGACCCCTCTCACATTTCTTCTCTGA 2080
Db 416 TCGCATTTGCTGAGCGGTGTGGTCACTGCCAAAGGAATGACCCCTCTCACATTTCTTCTCTGA 475
Qy 2081 TTGCGATACCCCGCGGCGCAGCTTGTCATCTCCCTCTTGGGCTTCCCGAGACACTAAGTCTG 2140
Db 476 TTGCGATACCCCGCGGCGCAGCTTGTCATCTCCCTCTTGGGCTTCCCGAGACACTAAGTCTG 535
Qy 2141 GAATGAAATTCACCTGCCTCTCAATTTGGCCACTGTGGGGGCGAGGGGTGACTTGGCT 2200
Db 536 GAATGAAATTCACCTGCCTCTCAATTTGGCCACTGTGGGGGCGAGGGGTGACTTGGCT 595
Qy 2201 TCCAGGCTGGAAGATTTATCTCACCCAGCCTAGCTATATATAAGGGCTGTGTGGAGGGG 2260
Db 596 TCCAGGCTGGAAGATTTATCTCACCCAGCCTAGCTATATATAAGGGCTGTGTGGAGGGG 655
Qy 2261 CTCACAGGCGCCAGTTCCAGGGTTTCATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAG 2317
Db 656 CTCACAGGCGCCAGTTCCAGGGTTTCATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAG 712

RESULT 7
BV076484 665 bp DNA linear STS 31-MAY-2003
LOCUS S212P60369PF7.T0 CZECHII/Ei Mus musculus STS genomic, sequence
DEFINITION
BV076484
ACCESSION
VERSION BV076484.1 GI:31192279
KEYWORDS tagged site.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 665)
Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,
Lander,E.S., Lindblad-Toh,K. and Daly,M.J.
```

TITLE The mosaic structure of variation in the laboratory mouse genome
JOURNAL Nature 420 (6915), 574-578 (2002)
MEDLINE 2234684
PUBMED 1246682
COMMENT

Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 665

Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SvimJ, C3H/HeJ, and BALB/cByJ. The WGS
reads were placed uniquely on the MGSCv3 C57BL/6J assembly and SNP
detection was carried out by SSNA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.

FEATURES

source

1..665
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="CZECHII/Ei"
/db_xref="taxon:10090"
/map="19 22-610 35751955-35751368"
/clone_lib="CZECHII/Ei"
<1..>665

STS
BASE COUNT 191 a 155 c 164 g 155 t
ORIGIN

Query Match 27.2%; Score 640.2; DB 11; Length 665;
Best Local Similarity .98.6%; Pred. No. 8.3e-176; Indels 1; Gaps 1;
Matches 656; Conservative 0; Mismatches 8;
QY 524 GTAGGAACCTGATGTATACCAAGGTCTGCTTCTTATGGTGGAGCCAGCAAGCATCGTGGGTG 583
DB 1 GTAGGAACCTGATGTATACCAAGGTCTGCTTCTTATGGTGGAGCCAGCAAGCATCGTGGGTG 60
QY 584 GAGCGAAGCGCAACCTCTCTAGCTCTGATCCATAGCAAGTAGGCTTAATGTTTCT 643
DB 61 GAGCGAAGCGCAACCTCTCTAGCTCTGATCCATAGCAAGTAGGCTTAATGTTTCT 120
QY 644 GTGCTAGGTGTTCATCTCTGTAATCGAGATCCTTGGCCCTTGCTTCAATTAGGGAGGCAC 703
DB 121 CTGCTAGGTGTTCATCTCTGTAATCGAGATCCTTGGCCCTTGCTTCAATTAGGGAGGCAC 180
QY 704 AAAATACTCAGAGATTCAAGACTGCTCAGCAGCCAGAGTCTTCTTCAAGGAAAGGTC 763
DB 181 AAAATACTCAGAGATTCAAGACTGCTCAGCAGCCAGAGTCTTCTTCAAGGAAAGGTC 240
QY 764 TCAACTCTCAGCCCCCTTAGCTCTGAGTCAGGCCCTGGAAACAAACCGCCACAGGAATGAG 823
DB 241 TCAACTCTCAGCCCCCTTAGCTCTGAGTCAGGCCCTGGAAACAAACCGCCACAGGAATGAG 300
QY 824 AAAAGCTGCCATAGCTGCTTGTACACTTCAAGAGGTCAAGAGAAATAGTTTACCAATGAA 883
DB 301 AAAAGCTGCCATAGCTGCTTGTACACTTCAAGAGGTCAAGAGAAATAGTTTACCAATGAA 360
QY 884 AACGAGAAGACCAACAGTTATCCATTGATAGCGTCTCAGGACAGATAGGACAGAGAAGC 943
DB 361 AACGAGAAGACCAACAGTTATCCATTGATAGCGTCTCAGGACAGATAGGACAGAGAAGC 420
QY 944 ACTAGAGAGGGGAACCCAGGAAGGACAAGGTATTAGTGTGTTGGTTTTCAGGGCAATGT 1003
DB 421 ACTAGAGAGGGGAACCCAGGAAGGACAAGGTATTAGTGTGTTGGTTTTCAGGGCAACGT 480
QY 1004 CTTGTACTGAAGATTCTAGAAACACAAATTTGCTGGTTGAACAGCTGAAGT-GGGGTGGG 1062

DB 481 CTCGTTCTGAAGGTTCTAGAAACAAATTTGCTGTTGAACAGCTGAAGTGGGGTGGG 540
QY 1063 GTTCTTACCCCATGTTTCATGGAAGGTGAGTGAGAGAGACAGATATATGATGGCCAGCA 1122
DB 541 GTTCTTACCCCATGTTTCATGGAAGGTGAGTGAGAGAGACAGATATATGATGGCCAGCA 600
QY 1123 TTAACAAACATACACAAACACCCCTAATTAACACCTTCCCTCTTCTACTGACACCCCTTCACT 1182
DB 601 TTAACAAACATACACAAACACCCCTAATTAACACCTTCCCTCTTCTACTGACACCCCTTCACT 660
QY 1183 CTCCT 1187
DB 661 CTCCT 665
AX468604 2074 bp DNA linear PAT 16-JUL-2002
Sequence 2 from Patent WO0246220.
ACCESSION AX468604
VERSION AX468604.1 GI:21901403
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Schwartz,B., Branellec,D. and Chien,K.
TITLE Sequences upstream of the carp gene, vectors containing them and
uses thereof
JOURNAL Patent: WO 0246220-A 2 13-JUN-2002;
Aventis Pharma S.A. (FR) ; The Regents of The University of
California at San Diego (US) ; Benoit, Patrick (FR)
FEATURES
source
1..2074
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 612 a 469 c 416 g 572 t 5 others
ORIGIN
Query Match 15.6%; Score 367.8; DB 6; Length 2074;
Best Local Similarity 76.4%; Pred. No. 5.9e-96;
Matches 597; Conservative 0; Mismatches 162; Indels 22; Gaps 11;
QY 1540 TCTCTGTGTCATCACTTCGCGCCGTTTGGGT--AGATCCTCTGATTAGCCTTCAGAT 1597
DB 1289 TCTATTCTTGACCACCTCTGATCCATTTTGAAGTAAATGCTCCAATTATTATGCTGTT 1348
QY 1598 TTAGAACACGGTGAGCCTGTGTGCACTAATTATGCGCAGTGACACCATAGATCAAGT 1657
DB 1349 TTAGAACACGGTAAGCATGTCATGTGCTA--ATGCGCAGTGACATCATAAAGAAAGT 1405
QY 1658 GCATTACTCAATGCTTTCATTTCTCTAATGCTGTCAGTGGCATGTGCACAGGSCCAT 1717
DB 1406 GCATTACTCAATGCTTTCATTTCTTAATGCTGTAAGTGGCATGTGCATGGGCCCTA 1465
QY 1718 TTTAGTGCAGACATCACTCCAGAGAAATTCAAACAGATAGAGACAAGTGGCACCCAGAC 1777
DB 1466 TTTAGC-CCAGACATCACTCCAAAGAAATTCAAACAGATATAGACAAGTGCCTTTAGGGC 1524
QY 1778 CCATCTCTCTCCCTCGGCTGATTTATCCCGAATAGGATGCTCCCAAGCAACACTTC 1837
DB 1525 CCAGATCCCTTCCCTTCAGGCTGTTTACCCAGGAATAGGATGCTCTGGGACAAAGTTTC 1584
QY 1838 CCAGCCAACTGGAGTCTGATAAGTCCAGTTATCAGAAAGATATGCTGTAAAGTGTGATG 1897
DB 1585 C---CTAAGTGAAGTGTGATAGTCTGCTTATCAGAAAGATATTACTGGGGGTGTGATA 1641
QY 1898 CACAGTGC--TTGCATTTTCTTGATACGTTAGTTCATATGAGAGCTGACAAAGAGGAAA 1955
DB 1642 TGTAGGGCATCTACATTTTCTTGATA-GGTAGTTCATATGAAGCTGACAAAGAA--AAAA 1698

QY 1956 AGACGAGGATGTGTCGAATATTAAACAGGAGCTGTCCCTGCTCCCGATACGTGGG 2015
 |||||
 Db 1699 AGGCGAGTGTGTTGTCGAATGTCAACAGACAGCTGTCCCTGTAC-TCTTGACAAATAGG 1757
 |||||
 QY 2016 ATGACTCGCATTTGCTGAGCGGTGGTCACTGCCAAAGGAATGACCCCTCTCACATTTCTT 2075
 |||||
 Db 1758 ATGACTTGCATTTGCTGAGCGATGTGATCACCAAAAGGAATGGCCCTCTCAATTTCTT 1817
 |||||
 QY 2076 CTTGATTCCGATACGCCCGCG- ----CCAGCTTGTTCATCTCCCTCTTTGGGCTTCCACAGAC 2130
 |||||
 Db 1818 CTTGATTACATATTTCAGCAGGGTTAGCTTGTCTCCCTCCCTCTTCAGCTTCCACAGAC 1877
 |||||
 QY 2131 ACTAAGTCTGGAATGAAATTCACCTGCTCTGAAATTTGGCACTGTGGGGGCGAGGGTG 2190
 |||||
 Db 1878 ACTGAGTCTGGAATGAAATTCACCTGCTCTGAGTTGGCTTCTTAATGGGGGGGGGAGTG 1937
 |||||
 QY 2191 TGACTTGGCTTCCAGAGCTGGAAGATTATCTACCCAGCGCTAGCTATATAA-CGGGCTG 2249
 |||||
 Db 1938 TTACTTCGGTTCAGAGGGTTAGCTTGTCTCCCTCCCTCTTCAGCTTCCACAGAC 1997
 |||||
 QY 2250 GTGTGAGGGGCTCCACAGGCGCAGTTCCAGGGGTTTCATCCACAGAGAGAGAAACATAG 2309
 |||||
 Db 1998 GTGTGAGGGGCGCCAGCAGGCGCAACTCCAGGGATTCTTC-CAOGACAGAAAACATAC 2056
 |||||
 QY 2310 A 2310
 |
 Db 2057 A 2057

RESULT 9

AF131884
 LOCUS 2074 bp DNA linear PRI 09-FEB-2000
 DEFINITION Homo sapiens cardiovascular-specific cardiac ankyrin repeat protein
 (CVARP) gene, 5'-flanking region and exon 1.
 ACCESSION AF131884
 VERSION AF131884.1 GI:6940841
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE Aihara,Y., Kurabayashi,M., Tanaka,T., Sekiguchi,K., Tomaru,K.,
 Kanai,H., Takeda,S. and Nagai,R.
 Human CVARP 5'-flanking region

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2074)

Aihara,Y.

TITLE Direct Submission

Submitted (26-FEB-1999) Second Department of Internal Medicine,
 Gunma University School of Medicine, 3-39-15 Showa-machi, Maebashi

371-8511, Japan

FEATURES Location/Qualifiers

1..2074
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

misc_feature

1..1832

/note="5'-flanking region"

1833..>2074

/gene="CVARP"

1833..>2074

/gene="CVARP"

/product="cardiovascular-specific cardiac ankyrin repeat

protein"

/note="CVARP"

1833..2074

/gene="CVARP"

/number=1

BASE COUNT 612 a 469 c 416 g 572 t

ORIGIN 5 others

Query Match 15.6%; Score 367.8; DB 9; Length 2074;
 Best Local Similarity 76.4%; Pred. No. 5.9e-96;
 Matches 597; Conservative 0; Mismatches 162; Indels 22; Gaps 11;

QY 1540 TCTCTCTGTGCATCACTTCGGCCCGTTTGGGGT--AGATCCCTCTGATTAGCCTTCAGAT 1597

|||

Db 1289 TCTATTTCTTGACCATCTCTGATCCATTTTGAAGTAAATGCTTCAATTTATATGCTGTT 1348

|||

QY 1598 TTAAACACAGGTGAGCCTGTGGTGCACATAATTATGGCCAGTGCACCATAGAGTCAAAGT 1657

|||

Db 1349 TTAAACACAGGTAAAGCATGTCAATGTGCTA---ATGGCCAGTGCATCATAAAGAAAAGT 1405

|||

QY 1658 GCATTTACTGAATGCTTTCAATTTCTCTAATGCTGTACGATGCATGTACAGGGCCAT 1717

|||

Db 1406 GCATTTACTGAATGCTTTCAATTTCTAATGCTGTAAAGTGGCATGTATGGGGCCTA 1465

|||

QY 1718 TTTAGTGTGCAGACATCACTCCAGAGAATTCCAAACAGATAGACAGTGCACCCAGAC 1777

|||

Db 1466 TTTAGC-CCAGACATCACTCCAAAGAAATTCCAAACAGATATAGACAAGTGCCTTTAGGGC 1524

|||

QY 1778 CCATCTCTTCCCTCGGGCTGAATTTATCCCGAATAATAGGATGTCCCAAGCAACACTTC 1837

|||

Db 1525 CCAGATCCCTTCCCTCAGGCTGTTTACCCAGGAATAGGATGTCTCGGCACAAGTTTCC 1584

|||

QY 1838 CCAGCCAACTGGAGTGTGATGAAGTCCAGTTATCAGAAAGATATGGCTGTAAAGTGTGATG 1897

|||

Db 1585 C---CTAAGTGAAGTGTGTATAGTCTGCTTATCAGAAAGATATTACTGGGGGTGTGATA 1641

|||

QY 1898 CACAGTGC--TTGCATTTCTTTGATACGTTAGTCATATGAGAGCTGACAAGAGGAAAA 1955

|||

Db 1642 TGTAGGGCATCTACATTTTCTTGATA--GGTAGTCATATGAAGCTGACAAAGAA--AAAA 1698

|||

QY 1956 AGACGAGCATGTGTGCAATATTAAACAGGACAGCTGTCCCTGGCTTCCCGATACGTGGG 2015

|||

Db 1699 AGGCGAGTGTGTGTGCAATGTCAACAGACAGCTGTCCCTGCAC-TCTTGACAAATAGG 1757

|||

QY 2016 ATGACTCGCATTTGCTGAGCGGTGTGTCTACTGCCAAAGGAATGACCCCTCTCACATTTCTT 2075

|||

Db 1758 ATGACTTGCATTTGCTGAGCGATGTGATCAACCAAAAGGAATGGCCCTCTCACATTTCTT 1817

|||

QY 2076 CCTGATTCCCATACGCCGCGG-----CCAGCTTGTTCATCTCCCTCTTGGGCTTCCACAGAC 2130

|||

Db 1818 CCTGATTTCATATTTCAGCAGGGTTAGCTTGTCTCCCTCTTTCAGCTTCCACAGAC 1877

|||

QY 2131 ACTAAGTCTGGAATGAAATTCACCTGCCTCTCAATTTGGCCACTGTGTGGGGGCGAGGGTG 2190

|||

Db 1878 ACTGAGTCTGGAATGAAATTCACCTGCCTCTGAGTTGGCTCTCTAATGGGGGCGGAGTG 1937

|||

QY 2191 TGACTTGGCTTCCAGGCTGGAAGATTATCTACCCAGCGCTAGCTATATAA-CGGGCTG 2249

|||

Db 1938 TTACTTTCGGTTCACAGGTTGGAAGATTATCTACCCGCGCCAGCTATATAAGCTGACCG 1997

|||

QY 2250 GTGTGAGGGGCTCCACAGGCGCAGTTCCAGGGGTTTCATCCCAAGAGAGAAAAACATAG 2309

|||

Db 1998 GTGTGAGGGGCGCCAGCAGGCGCAACTCCAGGGATTCTCTTC-CACGACAGAAAAACATAC 2056

|||

QY 2310 A 2310

|

Db 2057 A 2057

|||

RESULT 10

AL590622/c

LOCUS

DEFINITION

Human DNA sequence from clone Rpl1-320f15 on chromosome 10.

Contains the gene for ribonuclease P 30kD subunit, the gene for a

nuclear protein similar to CARP, ESTs, STSs and GSs, complete

sequence.

ACCESSION AL590622

VERSION AL590622.7

KEYWORDS HTG; CARP; ribonuclease P.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

AL590622 50111 bp DNA linear PRI 19-SEP-2001

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 50111)
Tracey.A.

Direct Submission
Submitted (19-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk

COMMENT

On May 31, 2001 this sequence version replaced gi:14161205.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr10>.
RP11-320F15 is from the library RPCI-11.2 constructed by the group of Fietzer de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone RP11-320F15 It may be shorter because we sequence overlapping sections only once, except for a short overlap
The true left end of clone RP11-236B18 is at 50012 in this sequence. The true right end of clone RP11-103A2 is at 100 in this sequence.

FEATURES

Source

1. 50111
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-320F15"
/clone_lib="RPCI-11.2"
complement(1..98)
/note="match: GSS: Em:AQ670367"
complement(1..97)
/note="match: GSS: Em:AQ544636"
complement(1..77)
/note="match: STS: Em:G56439
match: GSS: Em:AQ311708"
120..34469
/genes="BA320F15.1"
join(120..472,3273..3328,3962..4018,4428..4502,7467..7538,14216..14305,23179..23295,23858..23887,24284..24321,24717..24796,28974..29381)
/genes="BA320F15.1"
/product="BA320F15.1.1 (ribonuclease P (30kD) (RPP30))"
/notes="match: cDNAs: Em:U77665 Em:AK004137 Em:U95123
match: ESTs: Em:BE572689 Em:BE378859 Em:BE137893
Em:AA920575 Em:BG106820 Em:BF681617 Em:BF248001
Em:BF031745 Em:AA854455 Em:AI359795"
/evidence="not_experimental"
join(391..472,3273..3328,3962..4018,4428..4502,7467..7538,14216..14305,23179..23295,23858..23887,24284..24321,24717..24796,28974..29083)
/genes="BA320F15.1"
/note="match: proteins: Tr:P78346 Tr:O88796"

/codon_start=1
/evidence="not_experimental"
/product="BA320F15.1.1 (ribonuclease P (30kD) (RPP30))"
/protein_id="CAC70100.1"
/db_xref="GI:15717970"
/translation="MAVFADLDLRAGSLKALRGVTAHLGYSVVAINHIVDFKEK
KQIEKPVAVSELFITLPIVOQKSPILITRLTIIVSDPSHCNVLRTSSRARLYDV
VAVRPTEKLFHIACTHLDVLCITVTEKLPFVEKRPINVAIDGLAFELVYGP
KDSNRRVTISSALNMOICKGNVLISSAERPLEIRGPDVNLGLLPGLSSEDAK
AAVSTNCRAALLHGETRTARGLIISTVKKPSEDEDCPLPASKAKCEG"
689..800
/note="MIR repeat: matches 30..132 of consensus"
1443..1488
/note="L2 repeat: matches 2459..2502 of consensus"
1761..2155
/gene="BA320F15.1"
/note="match: GSS: Em:AQ805524"
1907..2434
/gene="BA320F15.1"
/note="match: GSS: Em:AQ775672"
1926..2208
/gene="BA320F15.1"
/note="match: GSS: Em:B92129"
2349..2482
/note="L2 repeat: matches 2557..2696 of consensus"
2939..3167
/note="MIR repeat: matches 13..262 of consensus"
complement(4176..4649)
/note="match: GSS: Em:AQ626500"
join(<4648..4764,7467..7538,14216..14305,23179..23887,
24284..24321,24717..24796,28974..29381)
/gene="BA320F15.1"
/product="BA320F15.1.2 (putative isoform 2)"
/note="match: ESTs: Em:AW939965"
/evidence="not_experimental"
4733..4858
/note="L2 repeat: matches 2601..2732 of consensus"
4928..5102
/note="L2 repeat: matches 2299..2478 of consensus"
5257..5439
/note="L2 repeat: matches 1808..1971 of consensus"
5440..5721
/note="AluX repeat: matches 7..293 of consensus"
5722..6237
/note="L2 repeat: matches 1200..1808 of consensus"
6884..7196
/note="AluJb repeat: matches 1..312 of consensus"
8243..8404
/note="2 copies 81 mer 81% conserved"
8445..8541
/note="HY1 repeat: matches 16..112 of consensus"
8542..8583
/note="HY1 repeat: matches 1..42 of consensus"
8850..9130
/note="AluX repeat: matches 1..289 of consensus"
9601..9741
/note="MIR repeat: matches 76..212 of consensus"
10798..11076
/note="AluY repeat: matches 1..291 of consensus"
11086..11219
/note="FLAM_C repeat: matches 1..129 of consensus"
11319..11491
/note="MERSA repeat: matches 3..189 of consensus"
11928..11984
/note="L2 repeat: matches 1743..1801 of consensus"
13755..13817
/note="L2 repeat: matches 2684..2749 of consensus"
15390..15671
/note="AluJO repeat: matches 1..282 of consensus"
15791..16088
/note="AluSg repeat: matches 1..296 of consensus"
16979..17036
/note="L2 repeat: matches 2423..2481 of consensus"

```
misc_feature      17236..17755
                  /gene="BA320F15.1"
repeat_region    /note="match: GSS: Em:AQ551877"
                  17709..17972
repeat_region    /note="AluX repeat: matches 45..304 of consensus"
                  19198..19363
repeat_region    /note="L1ME repeat: matches 5272..5435 of consensus"
                  19489..19928
repeat_region    /note="MLTIC repeat: matches 5..478 of consensus"
                  21838..22258
misc_feature     /note="I2 repeat: matches 2109..2669 of consensus"
                  22440..22946
                  /gene="BA320F15.1"
mrna             /note="match: GSS: Em:AQ676596"
                  join(23179..23295,23858..23887,24284..24321,24717..24795,
198974..29072,30647..30748,31596..31680,33706..34469)
                  /gene="BA320F15.1"
                  /product="BA320F15.1.3 (putative isoform 3)"
                  /note="match: ESTs: Em:BE567341 Em:A1292002 Em:BF684192
Em:BP435407 Em:A1740881 Em:AA651912 Em:BP438915"
                  /evidence=not experimental
                  24121..24518
misc_feature     /gene="BA320F15.1"
                  /note="match: GSS: Em:AQ026750"
                  complement(join(24284..24321,24724..24796,28974..29240))
misc_feature     /note="match: STS: Em:G24365"
                  24549..25020
                  /gene="BA320F15.1"
misc_feature     /note="match: GSS: Em:AQ223714"
                  24561..24956
                  /gene="BA320F15.1"
                  /note="match: GSS: Em:BS3566"
                  25321..25458
                  /note="FLAM C repeat: matches 1..133 of consensus"
misc_feature     complement(25417..25863)
                  /note="match: GSS: Em:AQ712354"
                  26204..26603
                  /note="L1M22 repeat: matches 5988..6302 of consensus"
                  28442..28516
                  /note="I2 repeat: matches 2636..2710 of consensus"
                  29149..29154
polyA_signal     /gene="BA320F15.1"
                  29166
polyA_site       /gene="BA320F15.1"
```

Query Match	14.7%;	Score 347.6;	DB 9;	Length 50111;	
Best Local Similarity	78.2%;	Pred. No. 7e-90;			
Matches 521;	Conservative	0;	Mismatches 129;	Indels 16;	Gaps 8;
Qy 1653	AAAGTCGATTACTGAATGCTTTCAAATTTCTCCTAAATGCTGGTACGATGGCATGTCACAGG	1712			
Db 50106	AAAGTCGATTACTGAATGCTTTCAAATTTCTTATATGATGGTAAAGTGGCATGTCATGGG	50047			
Qy 1713	GCCATTTTAGTCGACACATCACTCCAGAGAAATTCAAACAGATAGACAAAGTGGCAC	1772			
Db 50046	GCCTATTTAGCCCCACACATCACTCCAAAGAATTCCAAACAGATATAGCAAGTGCCTTT	49587			
Qy 1773	CAGACCCATCTCCTTCCCTCCGGCTGATATCCCCAGAAATAGGATGTCGCCAAGCAAC	1832			
Db 49986	AGGGCCACAGATCCCTTCCCTCAGGCTGTTTACCAGGGAATAGCATGTCC--TGGGAC	49930			
Qy 1833	ACTTCCAGGCAACTGGAGTGCTGATAGTCCAGTTATCAGAAAGATATGGCTGTAGTG	1892			
Db 49929	AACTTTTCCCCTAAGTGAAGTGTTGATAAGTCTGCTTATCAGAAAGATATTACTGGGGTG	49870			
Qy 1893	TGATGCACATGTC--TTGCATTTTCTTGATACGTTAGTTCATATAGAGCTGCACAAAGAG	1950			
Db 49869	TGATATGTATGGGCATCTACATTTTCTTGATA--GGTAGTCATATGAAAGCTGCACAAAGAA-	49812			
Qy 1951	GAATAAGACACCGATGTCGGTGCATATTAAACAGGCAGCTGTCCCTCGCTTCCCGATAC	2010			
Db 49811	-AAAAAGGGCAGTGATGTGGTGCATGTCAACAGACAGCTGTCCCTCGAC--TCTTGACAA	49754			

Qy	2011	GTGGATGACTCGCATTTGCTGAGCGGTGTGGTCACTGCCAAAGGAATGACCCCTCTCACAT	2070					
Db	49753	ATAGGATGACTTTGCATTGCTGAGCGGATGTGATCACCAACAAAGGAATGGCCCTCTCACAT	49694					
Qy	2071	TTCTTCTCGATTGGCATATAGCCGCGG-----CCAGCTTGTGCATCTCCCTCTTGGGCTTCC	2125					
Db	49693	TTCTTCTCGATTTCACATATTCAGCAGGGTTAGCTTGTCTCTCCCTCTCCCTCTTCACTTCC	49634					
Qy	2126	CAGACACTAAGTCTTGGAAATGAAAATTCACCTTGCTCTGAAATGGCCACTGGTGGGGCGAG	2185					
Db	49633	CAGACACTGAGTCTTGGAAATGAAAATTCACCTTGCTCTGAGTTGGCTCTTAATGGGGGTGG	49574					
Qy	2186	GGTGTGACTTGGCTTCCAGGCTGGAAGATTATCTACCCAGCCCTAGCTATATAA-CG	2244					
Db	49573	GAGTGTTACTTCCGTTCCAGGTTGGAAGATTATCTCACCGGCCCCAGCTATATAAGCT	49514					
Qy	2245	GCCTGCTGGAGGGGCTCCACAGGCGCCAGTTCCAGGGGTTCAATCACAAGAGAGAAAA	2304					
Db	49513	GACCGGTGTGAGGGGGCCAGCAGGCGCCAACTCCAGGGATTCTTTC-CACGACAGAAAA	49455					
Qy	2305	CATAGA 2310						
Db	49454	CATACA 49449						
RESULT 11								
LOCUS	AX322775	1901 bp	DNA linear PAT 07-JAN-2002					
DEFINITION	Sequence 19 from Patent WO0192567.							
ACCESSION	AX322775							
VERSION	AX322775.1	GI:18093755						
KEYWORDS								
SOURCE	unidentified							
ORGANISM	unclassified.							
REFERENCE	1							
AUTHORS	Bunk,D., Reuner,B., Beck,J. and Henkel,T.							
TITLE	Novel target genes for diseases of the heart							
JOURNAL	Patent: WO 0192567-A 19 06-DEC-2001;							
FEATURES	Medigene AG (DE)							
source	Location/Qualifiers							
	1..1901							
	/organism="unidentified"							
	/mol_type="genomic DNA"							
	/db_xref="taxon:32644"							
BASE COUNT	592 a 378 c 460 g 471 t							
ORIGIN								
Query Match 5.5%; Score 128.6; DB 6; Length 1901;								
Best Local Similarity 82.6%; Pred. No. 6.9e-26;								
Matches 171; Conservative 0; Mismatches 34; Indels 2; Gaps 2;								
Qy	2105	TCATCTCCCTCTTGGGCTTCCAGACACTAACTCTGGAATGAAAAATTCACCTGCCTCTGA	2164					
Db	27	TCCCTCCCTCTTACGTTCCAGACACTGATTCTGGAATGAAAAATTCACCTGCCTCTGA	86					
Qy	2165	ATTGSCCACTGGTGGGCGAGGGGTGACTTTGGCTTCCAGGCTCGAAGATTATCTCAC	2224					
Db	87	GTGGCTCTAATGGGGTGGAGTGTTACTTCGGTTCCAGGTTGGAAGATTATCTCAC	146					
Qy	2225	CCAGCCCTAGCTATATAA-CGGGCTGGTGTGAGGGGCTCCACAGGGCCAGTTCCAGGG	2283					
Db	147	CGGCCCCAGCTATATAAGCTGACCGGTGTGAGGGGCCACAGGCGCCAACTCCAGGA	206					
Qy	2284	TTCAATCCAAAGAGAGAAAAACATAGA	2310					
Db	207	TTCTTTC-CACGACAGAAAAACATACA	232					
RESULT 12								
LOCUS	BD094076	1901 bp	DNA linear PAT 27-AUG-2002					
DEFINITION	Shear stress-responsive DNAs.							

REFERENCE 1 (bases 1 to 1901)
 AUTHORS Myers,R.W.
 JOURNAL Unpublished (1996)
 COMMENT

Contact: Richard M. Myers
 Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine
 Department of Genetics, M-344, Stanford, CA 94305, USA
 Tel: 4157259687
 Fax: 4157259689
 Email: myers@shgc.stanford.edu

Primer A: GCATTTCGAGGCATGG
 Primer B: CCAGATGGATCATGAGG
 STS size: 222
 PCR Profile:

Initial incubation: 94 degrees C for 90 seconds
 Denaturation: 94 degrees C for 15 seconds
 Annealing: 62 degrees C for 23 seconds
 Polymerization: 72 degrees C for 30 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9600

Protocol:
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 Taq Polymerase: 0.05 units/ul
 Total Vol: 10 ul

Buffer: MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 20 mM
 pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from X83703
 -- Washington University/Merck EST sequence.

FEATURES

source 1..1901
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="10"
 STS 1246..1467
 primer_bind 1246..1263
 primer_bind complement(1447..1467)
 BASE COUNT 592 a 378 c 460 g 471 t
 ORIGIN

Query Match 5.5%; Score 128.6; DB 11; Length 1901;
 Best Local Similarity 82.6%; Pred. No. 6.9e-26;
 Matches 171; Conservative 0; Mismatches 34; Indels 2; Gaps 2;
 QY 2105 TCATCTCCCTTTGGGCTCCAGACACTAAGTCTGGAATGAAATTCACCTGCCCTCTGA 2164
 Db |||||||
 27 TCCCTCTCCCTTTTCAGCTCCAGACACTGATTCTGGAATGAAATTCACCTGCCCTCTGA 86
 QY |||||||
 2165 ATTGGCCATGTTGGGGCAGGGGTGACTTCGCTTCCAGGCTGGAGATTATCTCAC 2224
 Db |||||||
 87 GTTGGCTCTTAATGGGGGTGGGAGTGTACTTCGGTTCCAGGTTGGAAGATTATCTCAC 146
 QY |||||||
 2225 CCAGCCCTAGCTATATAA-CGGGCTGTGTGGAGGGGCTCCACAGGCGCCAGTTCACAGGGG 2283
 Db |||||||
 147 CCGGCCCCAGCTATATAAGTGTACCGGTGTGGAGGGGCCAGCAGGCGCCAACTCCAGGGA 206
 QY |||||||
 2284 TTCATCCACAAGAGAGAAAAACATAGA 2310
 Db |||||||
 207 TTCCTTC-CACGACAGAAAAACATACA 232

RESULT 15
 AC074094/c

LOCUS AC074094 160350 bp DNA linear HTG 01-SEP-2000
 DEFINITION Homo sapiens chromosome 10 clone RP11-236B18, WORKING DRAFT
 SEQUENCE, 17 unordered pieces.
 AC074094
 AC074094.3 GI:9958197
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 160350)
 AUTHORS Waterston,R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 160350)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (13-JUL-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 COMMENT On Sep 1, 2000 this sequence version replaced gi:9665205.
 ----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 ----- Project Information -----
 Center project name: H NH0236B18
 ----- Summary Statistics -----
 Sequencing vector: M13; 100%
 Sequencing vector: plasmid; 0%
 Chemistry: Dye-primer ET; 100% of reads
 Chemistry: Dye-terminator Big Dye; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 149545 bases at least Q40
 Consensus quality: 152532 bases at least Q30
 Consensus quality: 153924 bases at least Q20
 Insert size: 158750; sum-of-contigs
 Quality coverage: 4.85 in Q20 bases; agarose-fp
 Quality coverage: 4.71 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 17 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 2326: contig of 2326 bp in length
 * 2327 2426: gap of unknown length
 * 2427 7519: contig of 5093 bp in length
 * 7520 7619: gap of unknown length
 * 7620 15286: contig of 7667 bp in length
 * 15287 15386: gap of unknown length
 * 15387 23050: contig of 7664 bp in length
 * 23051 23150: gap of unknown length
 * 23151 33663: contig of 10513 bp in length
 * 33664 33763: gap of unknown length
 * 33764 50133: contig of 16370 bp in length
 * 50134 50233: gap of unknown length
 * 50234 68088: contig of 17855 bp in length
 * 68089 68188: gap of unknown length
 * 68189 88133: contig of 19945 bp in length
 * 88134 88234: gap of unknown length
 * 88234 113398: contig of 25164 bp in length
 * 113398 141732: gap of unknown length
 * 141732 141832: contig of 28235 bp in length
 * 141832 143187: gap of unknown length
 * 143187 143287: contig of 1355 bp in length
 * 143287 143288: gap of unknown length
 * 143288 144942: contig of 1655 bp in length

* 144943 145042: gap of unknown length
* 145043 147286: contig of 2244 bp in length
* 147287 147386: gap of unknown length
* 147387 150852: contig of 3466 bp in length
* 150853 150952: gap of unknown length
* 150953 153550: contig of 2598 bp in length
* 153551 153650: gap of unknown length
* 153651 157257: contig of 3607 bp in length
* 157258 157358: gap of unknown length
* 157358 160350: contig of 2993 bp in length.

FEATURES

Location/Qualifiers

1. .160350

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="10"

/clone="RP11-236B18"

1. .2326

/note="assembly_name:Contig10"

misc_feature

2427. .7519

misc_feature

/note="assembly_name:Contig11"

misc_feature

7620. .15286

misc_feature

/note="assembly_name:Contig12"

misc_feature

15387. .23050

misc_feature

/note="assembly_name:Contig13"

misc_feature

23151. .33663

misc_feature

/note="assembly_name:Contig14"

misc_feature

33764. .50133

misc_feature

/note="assembly_name:Contig15"

misc_feature

50234. .68088

misc_feature

/note="assembly_name:Contig16"

misc_feature

68185. .88133

misc_feature

/note="assembly_name:Contig17"

misc_feature

88234. .113397

misc_feature

/note="assembly_name:Contig18"

misc_feature

113498. .141732

misc_feature

/note="assembly_name:Contig19"

misc_feature

141833. .143187

misc_feature

/note="assembly_name:Contig3"

misc_feature

143288. .144942

misc_feature

/note="assembly_name:Contig4"

misc_feature

145043. .147286

misc_feature

/note="assembly_name:Contig5"

misc_feature

147387. .150852

misc_feature

/note="assembly_name:Contig6"

misc_feature

150953. .153550

misc_feature

/note="assembly_name:Contig7"

misc_feature

153651. .157257

misc_feature

/note="assembly_name:Contig8"

misc_feature

157358. .160350

misc_feature

/note="assembly_name:Contig9"

BASE COUNT 47390 a 32724 c 33135 g 45498 t 1603 others
ORIGIN

Query Match 5.4%; Score 127.2; DB 2; Length 160350;

Besc Local Similarity 52.5%; Pred. No. 2.9e-25; Indels 2; Gaps 1;
Matches 171; Conservative 0; Mismatches 153;

QY 1540 TCCTCTGTGCATCACTTCGGCCGCTTTGGGGT--AGATCCTCTGATTAGCCTTCAGAT 1597
DB 113707 TCTATTCTTGACCATCTGATCCATTTTGAAGTAAATGCTCCAATTATTATGCTGT 113648
QY 1598 TTAGAACACGGTAGCCTGTGGTGCACTAATTATGGCCAGTGACACCATAGACTCAAAGT 1657
DB 113647 TTAGAACACGGTAAGCATGTGCTGCTAATTATGGCCAGTGACATCAATAAAGAAAGT 113588
QY 1658 GCATTACTGAATGCTTTCAATTTCTCCTAATGCTGATGATGATGATGATGATGATGATGAT 1717
DB 113587 GCATTACTGAATGCTTTCAATTTCTTATAATGATGATGATGATGATGATGATGATGAT 113528
QY 1718 TTTAGCTGCAGACATCACTCCAGAGAAATCCAAACAGATAGACAAAGTGGCACCCAGAC 1777
DB 113527 TTTAGCCCCAGACATCACTCCAAAGAAATTCNNNNNNNNNNNNNNNNNNNNNNNNNN 113468

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 03:46:02 ; Search time 608.653 Seconds
(without alignments)
10457.963 Million cell updates/sec

Title: US-10-005-337A-1
Perfect score: 2358
Sequence: 1 99atccttcatgtttaaca.....caggctggagccaccatg 2358

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2358	100.0	2358	24	ABV73020 Mouse CARP protein
2	1763.4	74.8	2247	21	AAA10406 Murine cardiac ank
3	367.8	15.6	2074	24	ABV73021 Human CARP protein
4	130.2	5.5	1988	22	AA193508 Human polynucleoti
5	128.6	5.5	1901	22	AAH02910 Human shear stress
6	128.6	5.5	1901	24	AAH02910 Human CAA56676 pro
7	95	4.0	7280	25	ABZ24607 Mouse gastric antr
8	95	4.0	7280	25	AAZ50815 Mouse pre-AMP-18 g

C	9	89.2	3.8	18105	25	ABZ82337	Mouse histamine re
C	10	88.8	3.8	659158	25	ABX16390	Mouse high growth
C	11	88.6	3.8	48974	20	AAZ55300	Mouse Presenilin-1
C	12	83.6	3.5	707	24	ABT09109	Phase-1 Rat CT gen
C	13	82	3.5	659158	25	ABX16390	Mouse high growth
C	14	80.4	3.4	10917	24	ABA97708	Wild-type mouse Wo
C	15	78	3.3	892	18	AAV06155	Viral infection ge
C	16	76.8	3.3	249487	24	ABN85733	Mouse genomic regi
C	17	76.6	3.2	5023	24	ABL56461	Nucleotide sequenc
C	18	76.6	3.2	17758	24	ABL56463	Alphaf locus of al
C	19	75.8	3.2	335	24	ABZ69028	Novel murine polyn
C	20	75.6	3.2	5990	24	ABK52570	Rat sequence diffe
C	21	75.6	3.2	7280	25	ABZ24607	Mouse gastric antr
C	22	75.6	3.2	7280	25	AAZ50815	Mouse pre-AMP-18 g
C	23	75	3.2	49999	20	AAZ23891	Murine LOBO genom
C	24	75	3.2	49999	20	AAZ23896	Murine LOBO genom
C	25	74.8	3.2	37940	20	AAZ01026	Partial mouse PGL
C	26	74.6	3.2	10614	16	AAQ89555	Hamster cholestero
C	27	73.4	3.1	4164	19	AAV12216	Mouse retinoid met
C	28	73.4	3.1	4164	24	AAZ24512	Mouse P450RAI geno
C	29	71.6	3.0	171936	24	ABZ56565	Human SULF2 genom
C	30	69.4	2.9	1889	24	AAZ94903	Human DNA sequence
C	31	69.4	2.9	1889	25	ABZ77648	Differentially exp
C	32	68.6	2.9	4072	22	AAZ85818	Murine sequence fr
C	33	68	2.9	1094	20	AAZ57479	Rat U3 gene trap d
C	34	68	2.9	16442	18	AAZ83006	Partial mouse WRN
C	35	65.6	2.8	29604	18	AAZ83005	Partial mouse WRN
C	36	65	2.8	6645	17	AAZ65002	Mouse cell cycle r
C	37	63	2.7	874	20	AAZ57448	Rat U3 gene trap d
C	38	63	2.7	874	20	AAZ57470	Rat U3 gene trap d
C	39	63	2.7	3240	22	AAZ05314	Mouse alpha-1,3 ga
C	40	63	2.7	3537	22	AAZ05315	Mouse alpha-1,3 ga
C	41	62	2.6	7208	21	AAA40866	Murine tumour necr
C	42	61.6	2.6	8402	24	ABZ87055	Genomic DNA encodi
C	43	60.6	2.6	29392	19	AAZ15422	Mouse poly Ig rece
C	44	60.4	2.6	696	24	ABQ96847	Mouse ES cell rela
C	45	59.8	2.5	347	24	ABQ97071	Mouse ES cell rela

ALIGNMENTS

RESULT 1

ABV73020
ID ABV73020 standard; DNA; 2358 BP.

XX AC ABV73020;

XX DT 08-JAN-2003 (first entry)

XX XX Mouse CARP protein coding sequence upstream DNA fragment.

XX DE Cardiac ankyrin repeat protein; CARP; cardiant; immunosuppressive;
XX KW antiinflammatory; gene therapy; antisense gene therapy; mouse; ds.

XX OS Mus musculus.

XX PN WO200246220-A2.

XX XX 13-JUN-2002.

XX PF 05-DEC-2001; 2001WO-EP15412.

XX PR 07-DEC-2000; 2000US-251582P.

XX XX (AVET) AVENTIS PHARMA SA.
(REGC) UNIV CALIFORNIA.
(BENO/) BENOIT P.

XX PI Schwartz B, Branellec D, Chien K;
XX DR WPI; 2002-740642/80.

Db 1621 GCACCTAATTATGGCCAGTGCACCATAGATGATCAAGTGCATTACTGATGCTTCAATTT 1680
Qy 1681 CTCCTAATGCTGGTACGATGCATGTCAAGGGCCATTTTAGCTGCAGACATCACTCCAG 1740
Db 1681 CTCCTAATGCTGGTACGATGCATGTCAAGGGCCATTTTAGCTGCAGACATCACTCCAG 1740
Qy 1741 AGAATTCACACAGATAGACACAGTGGCACCAGCCATCTCTTCCCTCGGGTGA 1800
Db 1741 AGAATTCACACAGATAGACACAGTGGCACCAGCCATCTCTTCCCTCGGGTGA 1800
Qy 1801 TTATCCCCAGAAATAGGATGTCACCAAGACACACTTCCAGCCAACTGGAGTGTGATAA 1860
Db 1801 TTATCCCCAGAAATAGGATGTCACCAAGACACACTTCCAGCCAACTGGAGTGTGATAA 1860
Qy 1861 GTCAGTTATCAGAAAGATATGGCTGTAAAGTGTGATGCACAGTCTTGCATTTCTTGTAT 1920
Db 1861 GTCAGTTATCAGAAAGATATGGCTGTAAAGTGTGATGCACAGTCTTGCATTTCTTGTAT 1920
Qy 1921 ACCTTAGTCATATCAGAGCTGACAAAGAGGAAAAAGAGCAGCGATGTGTGCAATATTA 1980
Db 1921 ACCTTAGTCATATCAGAGCTGACAAAGAGGAAAAAGAGCAGCGATGTGTGCAATATTA 1980
Qy 1981 ACAGGAGCTGTCCCTCGCTTCCCGATACCTGGGATGACTCGCATTTCTGAGCGGTGTG 2040
Db 1981 ACAGGAGCTGTCCCTCGCTTCCCGATACCTGGGATGACTCGCATTTCTGAGCGGTGTG 2040
Qy 2041 GTCATGCCAAAGGAATGACCTCTCACATTTCTTCTGATTCGCATACCGCGGCCAG 2100
Db 2041 GTCATGCCAAAGGAATGACCTCTCACATTTCTTCTGATTCGCATACCGCGGCCAG 2100
Qy 2101 CTTGTGATCTCCCTTGGGCTTCCAGACACATTAAGTCTGGAATGAAAAATTCACCTGCCT 2160
Db 2101 CTTGTGATCTCCCTTGGGCTTCCAGACACATTAAGTCTGGAATGAAAAATTCACCTGCCT 2160
Qy 2161 CTGAATTTGGCCACTGTGTGGGGCAGGGGTGTGACTTGGCTTCCAGGCTGGAAGATTATC 2220
Db 2161 CTGAATTTGGCCACTGTGTGGGGCAGGGGTGTGACTTGGCTTCCAGGCTGGAAGATTATC 2220
Qy 2221 TCACCCAGCCCTAGCTATATAACGGGCTGTGTGGAGGGCTCCACAGGGCCAGTTCCAG 2280
Db 2221 TCACCCAGCCCTAGCTATATAACGGGCTGTGTGGAGGGCTCCACAGGGCCAGTTCCAG 2280
Qy 2281 GGGTTTCATCACAAGAGAGAAAAACATAGACTCGAGGTCTAGGAGCTTGCATGCTGCA 2340
Db 2281 GGGTTTCATCACAAGAGAGAAAAACATAGACTCGAGGTCTAGGAGCTTGCATGCTGCA 2340
Qy 2341 GGTCCGAGGCCCATGG 2358
Db 2341 GGTCCGAGGCCCATGG 2358

RESULT 2

AA10406
ID AA10406 standard; DNA; 2247 BP.
XX AC
XX AC
XX AA10406;
DT 18-JUL-2000 (first entry)
XX DE
XX DE Murine cardiac ankyrin repeat protein (CARP) promoter.
XX DE Cardiac ankyrin repeat protein; CARP promoter; murine; adenovirus vector;
KW cardiac specific; heart disease; gene therapy; ds.
XX OS
XX OS Mus musculus.
XX PN
XX PN W0200015821-A1.
XX PD
XX PD 23-MAR-2000.
XX PF
XX PF 10-SEP-1999; 99WO-US20730.
XX PR
XX PR 11-SEP-1998; 98US-0099960.

XX (REGC) UNIV CALIFORNIA.

XX Chien KR, Wang Y, Evans S;

XX WPI; 2000-271457/23.

XX Human type-5 recombinant adenovirus vector used for targeted gene
PT therapy for heart disease and evaluating gene function contains a
PT tissue-restricted promoter and inverted terminal repeat sequences -

XX Claim 8; Page 29-30; 33pp; English.

XX The invention relates to a human type-5 recombinant adenovirus vector
CC for achieving cardiac-restricted transcription of a gene of interest.
CC The vector comprises inverted terminal repeat (ITR) sequences from human
CC adeno-associated virus (AAV) type 2 (AA10404-A10405) and a cardiac
CC tissue-specific promoter. In particular, the promoter is that of the
CC cardiomyocyte-restricted cardiac ankyrin repeat protein (CARP) gene.
CC The adenovirus vector is used for targeted gene therapy for heart
CC disease and for evaluating gene function. Cardiac restricted
CC transcription of a transgene in both neonatal and mature cardiac tissues
CC can be achieved to treat inherited and acquired heart diseases. The
CC vector is suitable for tissue-specific use in vivo and in vitro and
CC provides cardiac restricted transcription. The present sequence
CC represents the murine cardiac ankyrin repeat protein (CARP) promoter.

XX Sequence 2247 BP; 609 A; 549 C; 485 G; 589 T; 15 other;

Query Match 74.8%; Score 1763.4; DB 21; Length 2247;
Best Local Similarity 91.2%; Pred. No. 0;
Matches 2124; Conservative 0; Mismatches 69; Indels 137; Gaps 19;

Qy 1 GGATCTCTTTCATGTTAAACAATATCAACCTAACCCAGGGGACACGCCCTGCCTGACAGT 60

Db 31 GGATCTCTTTCATGTTAAACAATATCAACCTAACCCAGGGGACACGCCCTGCCTGACAGT 90

Qy 61 GGCTTTGCCACCCATGAATACTTCTAGTCTAGTCCGTTTGTGAAATCTCAGCCCATCCCA 120

Db 91 GGCTTTGCCACCCATGAATACTTCTAGTCTAGTCCGTTTGTGAAATCTCAGCCCATCCCA 150

Qy 121 ACATCTTGTGCAAGCCCATCTCTACAGGTGCTCATTTGGGAATTTCTCGAGGTTCTCT 180

Db 151 ACATCTTGTGCAAGCCCATCTCTACAGGTGCTCATTTGGGAATTTCTCGAGGTTCTCT 210

Qy 181 TTCAGGATCAGCTGATTTCTAGGCGACGCTTCTACACCTGGGGCTTCACCCCTTTGG 240

Db 211 TTCAGGATCAGCTGATTTCTAGGCGACGCTTCTACACCTGGGGCTTCACCCCTTTGG 270

Qy 241 GGGATCAAAACGACCCCTTTACAGGGGTACATATCATATCTATATGTCAGGTATTTA 300

Db 271 GGGATCAAAACGACCCCTTTACAGGGGTACATATCATATCTATATGTCAGGTATTTA 330

Qy 301 CATTAGCATTCGTAACAGTAGCAAAATTAAGGTATGAATAGCAATGAATAATTTTAT 360

Db 331 CATTAGCATTCGTAACAGTAGCAAAATTAAGGTATGAATAGCAATGAATAATTTTAT 390

Qy 361 GATTGAAGTCCACCAACATGAGCGCCACACTGTTCTAGAGAAAAATCAGCTGGGTG 420

Db 391 GATTGAAGTCCACCAACATGAGCGCCACACTGTTCTAGAGAAAAATCAGCTGGGTG 450

Qy 421 GGGAAAGGTTTGGAAAGCCCTTTCTGTCCATTCTTCAATCTTCAAGTGTATGTTTCA 480

Db 451 GGGAAAGGTTTGGAAAGCCCTTTCTGTCCATTCTTCAATCTTCAAGTGTATGTTTCA 510

Qy 481 GAAAGCCTTTCAGCTGTTCTGTGGGCTTTAGTAAGTCTAGTAGGAATCTGTATGTAC 540

Db 511 GAAAGCCTTTCAGCTGTTCTGTGGGCTTTAGTAAGTCTAGTAGGAATCTGTATGTAC 570

Qy 541 CAGGTCTGCTTCTTATGGGTGGAGCCCAAGCAGCATGTTGGGTGGAGGAAGACGACCT 600

Db 571 CAGGTCTGCTTCTTATGGGTGGAGCCCAAGCAGCATGTTGGGTGGAGGAAGACGACCT 630

Qy		601	CACCTTCTAGCTCTGCAATCCATAGCAAGTAGCCTAATGTTCCTTGCTCTAGGTGTCATCT	660
Db		631	CACCTTTCTA-CTCTGCATCCATAGCAAGTAGCCTAATG-TTCTGNGTCTAGG-GTCACTCT	687
Qy		661	CTGTGAATCAGATCCCTTGG-CTTGTGCTTCAAATTAGGAGGCCAAAACTACTCAGAGATT	719
Db		688	CTGTGAATCAGAGATCCCTTGGCCCTTGTGTGAATTAGGAGGCCAAAACTTTAAAATAAT	747
Qy		720	CAAGACTGCTCAGCAGCCCCAGA--GTCTTTCCTCAAGAAAGGTCTCAACTCTCAGCCC	777
Db		748	CAAGACTGNTCAACANNCANAAGTCCTTCTCARRAAGGAAGNCTTAACNTWNAACC	807
Qy		778	CCCTTAGCTCT--GAGTCAGCCTCGGAAC--AAACGGCCACAGAAATAGTGTAAACCATGAAACG	829
Db		808	CCCTTTACTTTTGAGTCAAGCCTCGGAACAAACCCGGCCCCAGGAATGAAAAAGCTTGC	867
Qy		830	--TGCCATAGCTGCTGTCTACTTCAAGAGGTCAAGAAATAGTGTAAACCATGAAACG	887
Db		868	CATNACCTGGTGGCCCTTTTNAANAGNCAAAAAAAAAATTTGTGTTAACNTTTGAAAAAC	927
Qy		888	AGAAGACCAACAGTTATCCATTGATAGCGTCTCAGGACAGATAGGACAGAGAGAACACTA	947
Db		928	CGAAGACCAACAGTTATCC-----	946
Qy		948	GGAGAGGGGAACCCACGAAGGACAAGGTATTAGTGTGTGTTTTCAGGGCAATGTCTTG	1007
Db		947	-----	946
Qy		1008	TACTGAAGATTCTAGAAAACAATAATTTGCTGGTTGAACAGCTGAAGTGGGTGGGGTTCT	1067
Db		947	-----TCTAGAAAACAATAATTTGCTGGTTGAACAGCTGAAGTGGGTGGGGTTCT	996
Qy		1068	TACCCCATGTTTCATGGAAGGAGTGAAGTGGAGAGACAGATATATGATGGCCAGCATAACA	1127
Db		997	TACCCCATGTTTCATGGAAGGAGTGAAGTGGAGAGACAGATATATGA--GGCCAGCATACA	1055
Qy		1128	AACATACACAACACCCCTAAATTAAACATTCCTCTTTACTGACACCCCTTCTCACTCTCT	1187
Db		1056	AACATACACAACACCCCTAAATTAAACATTCCTCTTTACTGACACCCCTTCTCACTCTCT	1115
Qy		1188	CTTTCATAAAAAATAAAAAAGTATTTATGTGGCTCTTACGATAGATCTTTCTCTCGAA	1247
Db		1116	CTTTCATAAAAAATAAAAAAGTATTTTA--GTGGCTCTTACGATAGAACTTTTCTCTCGAA	1174
Qy		1248	CTATAAAAAGATCTAAATATTTATTTTTCACATTTTATATCTTAGCCATGACAAAGCC	1307
Db		1175	CTATAAAAAGATCTAAATATTTATTTTTCACATTTTATATCTTAGCCATGACAAAGCC	1234
Qy		1308	AGAAACAAGTATTTTTTGGCCTCTCTCAACAGCAAAAGCTTGGGGCCTTTTTTGTTCCTGT	1367
Db		1235	AGAAACAAG-ATTTTTTGGCCTCTCTCAACAGCAAAAGCTTGGGGCCTTTTTTGTTCCTGT	1293
Qy		1368	TAGGAATAGAACACGAGAGCCCGGTGATCTAGGCAGATGCTCTATCATTTAGCCCATGAG	1427
Db		1294	TAGGAATAGAACACGAGAGCCCGGTGATCTAGGCAGATGCTCTATCATTTAGCCCATGAG	1353
Qy		1428	TCTCCAGCCTCAGACGCACATTTTCTCGGGCTCTTTAGACTTTTCCACAGATTTGG	1487
Db		1354	TCTCCAGCCTCAGACGCACATTTTCTCGGGCTCTTTTAGACTTTTCCACAGATTTGG	1413
Qy		1488	AAACTTTACTGACAGCATCCAAGTTGTGCTTCTGCTAAGAACTGGACTCACATCTCTCTG	1547
Db		1414	AAACTTTACTGACAGCATCCAAGTTGTGCTTCTGCTAAGAACTGGACTCACATCTCTCTG	1473
Qy		1548	TGCATCACTTGGCCCGTTTTTGGGGTAGATCTCTTGATTAGCTTTCAGATTTTAAACACG	1607
Db		1474	-GCATCACTTGGCCCGTTTTTGGGGTAGATCTCTTGATTAGCTTTCAGATTTTAAACACG	1532
Qy		1608	GTGAGCCTGTGGTCCACTTAATTATGGCCAGTGACACCATAGATGCAAGTGCATTTACTGA	1667
Db		1533	GTGAGCCTGTGGT--CACTTAATTATGGCCAGTGACACCATAGATGCAAGTGCATTTACTGA	1591
Qy		1568	ATGCTTTCAATTTCTCTTAATGCTGGTACGATGTCATGTGCACAGGGCATTTTAGCTGCA	1727

1592	ATGCTTTCAATTTCTCTTAATGCTGGTACGATGGCATGTCCAGGGCCATTTTAGCTGCA	1651
1728	GACATCACTCCAGAGAAATTCCAAACAGATAGACAAAGTGGCCACCCAGACCCCATCTCCTT	1787
1652	GACATCA - TCCAGAGAAATTCCAAACAGATAG - GACAAGTGGCACCAGACCCCATCTCCTT	1709
1788	CCCTCGGGCTGATTATCCCCAGAAATAGAGTGTCCAAAGCAACACTTCCCAGCCAAC	1847
1710	CCCTCGGGCTGATTATTCGCCA - AAATAGAGTGTCCAAAGCAACACTTCCCAGCCAAC	1768
1848	GGAGTGTGATAAGTCCAGTTCATCAGAAAAGATATGGCTGTAAGTGTGATGCACAGTGC	1907
1769	GGAGTGTGATAAGTCCAGTTCATCAGAAAAGATATGGCTGTAAGTGTGATGCACAGTGC	1828
1908	GCATTTCTTGTATAGTTAGTTCATATGAGAGCTGACAAAGAGAAAAGAGCAGCGATG	1967
1829	GCATTTCTTGTATAGTTAGTTCATATGAGAGCTGACAAAGAGAAAAGAGCAGCGATG	1888
1968	TGGTGCAAATATTAAACAGGAGCTGTCCCTTGGCTTCCCGATACGTGGAGTGACTCCGATT	2027
1889	T - GTGCAATATTAAACAGGAGCTGTCCCTTGGCTTCCCGATACGTGGAGTGACTCCGATT	1947
2028	GCTGAGCGGTGTGGTCACTGCCAAAGAAATGAACCTCTCACATTTCTTCTGATTCGCAT	2087
1948	GCTGAGCGGTGTGGTCACTGCCAAAGAAATGACCCCTCTCATTTCTTCTGATTCGCAT	2007
2088	ACGCCGGCGGCAGCTTGTCACTCTCCCTTTGGGCTTCCAGACACTAAAGTCTGGAATGAA	2147
2008	ACGCCGGCGGCAGCTTGTCACTCTCCCTTTGGGCTTCCAGACACTAAAGTCTGGAATGAA	2067
2148	AAATCACTGCCTCTGAATTTGGCACTGTGTGGGGCAGGGGTGTGACTTTGGCTTCCGAG	2207
2068	AAATCACTGCCTCTGAATTTGGCACTGTGTGGGGCAGGGGTGTGACTTTGGCTTCCGAG	2127
2208	CTGGAAGATTATCTCACCCAGCCCTAGCTATATAACGGGCTGGTGTGAGGGGCTCCACA	2267
2128	CTGGAAGATTATCTCACCCAGCCCTA - CTATATACGGGCTGGTGTGAGGGGCTCCACA	2186
2268	GGGCCAGTTCACGGGGTTCATCCACAAGAGAGAAAAACATAGACTCCAGG	2317
2187	GGGCCAGTTCACGGGGTTCATCCACAAGAGAGAAAAACATAGACTCCAGG	2236

RESULT 3
ABV73021
ID ABV73021 standard: DNA: 2074 BP.

AC ABV73021:

XX DT 08-JAN-2003 (first entry)

XX
DE Human CAPP protein coding sequence upstream DNA fragment.

Cardiac ankyrin repeat protein; CARP; cardiant; immunosuppressive;
xx
kw
kw
kw
anti-inflammatory; gene therapy; antisense gene therapy; human; ds.
xx

XX.
Us
Homo sapiens.

XX PN WQ200246220-A2.

13-JUN-2002
XX PDXX
PF 05-DEC-2001: 2001WO-EP15412.XX
PR 07-DEC-2000: 2000US-251582P.XX
PA (AVET) AVENTIS PHARMA SA.

PA (REGC-) UNIV CALIFORNIA.
PA (BENO/) BENOIT P.

XX
PI Schwartz B. Brar

XX

DR WPI; 2002-740642/80.

XX New promoter sequence derived from a portion upstream of the coding
PT sequence of a gene for Cardiac Ankyrin Repeat Protein, for controlling
PT the level and specificity of expression of a transgene in cardiac
PT muscle cells.

XX
XX
PS Claim 5; Fig 2; 48pp; English.

XX The invention relates to a polynucleotide (I) comprising a fragment of a
CC sequence upstream of the coding part of the gene for the Cardiac Ankyrin
CC Repeat Protein (CARP). (I) is capable of inducing a specific expression
CC in vivo of a gene operably linked to (I), in cardiac cells. (I) or a
CC vector (Iib) comprising (I) is useful for the manufacture of a medicament
CC intended for the treatment of cardiac insufficiency, cardiac hypertrophy
CC and hypoxia, and for preventing rejection during cardiac transplant. An
CC expression cassette under the control of (I) is useful for encoding a
CC protein or RNA which is capable of activating the growth of cardiac
CC cells, reducing or suppressing an immune response, inducing angiogenesis,
CC correcting muscle contractility, cardiac hypertrophy, cardiac
CC insufficiency and myocarditis. (Iib) is useful for expressing a gene of
CC therapeutic interest in vivo, by isolating (Iib) and introducing (Iib) in
CC the cardiac tissue, under conditions so that the gene of interest is
CC expressed. (I), the vectors and the compositions are useful in clinical,
CC experimental, therapeutic and diagnostic fields, and in the treatment and
CC prevention of cardiac pathologies. (I) is also useful for generating
CC transgenic animals which constitute models for studying certain cardiac
CC pathologies. The transgenic animals are also useful for screening
CC molecules for their activity on the regulatory sequences of the gene
CC encoding the CARP protein. The present sequence represents the DNA
CC fragment upstream of the coding sequence of a human CARP protein.

XX
SQ Sequence 2074 BP; 612 A; 469 C; 416 G; 572 T; 5 other;

Query Match 15.6%; Score 367.8; DB 24; Length 2074;
Best Local Similarity 76.4%; Pred. No. 5.9e-97;
Matches 597; Conservative 0; Mismatches 162; Indels 22; Gaps 11;

QY 1540 TCTCTCTGTCATCCTTCGCGCCGCTTTTGGGGT--AGATCCTCTGATTAGCCTTCAGAT 1597
DB TCTATTCTTTGACCACTCTGATCCATTTTGAAGTAAATGCTCCAATTATTATGCTGTT 1348
QY 1598 TTAGAACACGGTGAGCGCTGGTGCACTAATTATGGCCAGTGACACCATAGAGTCAAAGT 1657
DB TTAGAACACGGTAAGCATGTGTCATGCTGA---ATGGCCAGTGACATCATAAAGAAAAGT 1405
QY 1658 GCATTACTGAATGCTTTCATTTCTCTAATGCTGTGATGCTGACATGTCACAGGGCCAT 1717
DB GCATTACTGAATGCTTTCATTTCTCTAATGCTGTGATGCTGACATGTCACAGGGCCAT 1465
QY 1718 TTTAGTTCAGACATCACTCCAGAGAAATCCAAACAGATAGACAAAGTGGCACCCAGAC 1777
DB TTTAGC-CCAGACATCACTCCAAAGAAATCCAAACAGATATAGACAAAGTGCCTTTAGGGC 1524
QY 1778 CCAATCTCTTCCCTCGGGCTGATTAATCCCAAGAAATAGGATGTCCCAAGCAACACTTC 1837
DB CCAGATCCCTTCCCTCCAGCGCTGTTTACCCAGGGAATAGGATGCTCTGGGACAAAGTTTC 1584
QY 1838 CCAGCCCACTGGAGTGTGATAGTCCAGTTATCAGAAAGATATGCTGTAAAGTGTATG 1897
DB C---CTAAGTGAAGTGTGATAGTGTGCTTATCAGAAAGATATTAATGGGGGTGATA 1641
QY 1898 CACAGTGC--TTGCATTTCTTGATAGTGTAGTATATAGAGTGCACAAAGAAAGAAA 1955
DB TGTAGGCGCATCTACATTTCTTGATA-GGTAGTATATGAAGCTGACAAAGAA--AAAA 1598
QY 1956 AGAGCAGCGATGTGGTCAATATTAACAGGACAGTGTCCCTGGCTTCCCGATACGTGGG 2015
DB AGGGCAGTGTGTTGGTCAATGTCAACAGACAGCTGTCCCTGCAC-TCTTGCACAAATAGG 1757
QY 2016 ATGACTCGCATTCCTGAGCGGTGTGTCATGTCACCAAGGAATGACCTCTCAATTTCTT 2075
DB ATGACTTGCATTCCTGAGCGATGTGATCAACCAAGGAATGGCCCTCTCAATTTCTT 1817

QY 2076 CTTGATTTCGATACAGCGCGG-----CCAGCTTGTTCATCTCCCTCTTGGCTTCCACAGAC 2130
DB CTTGATTTCATATTCAGCAGCGGTAGCTGTCTCCCTCTCCCTCTTTCAGCTTCCACAGAC 1877
QY 2131 ACTAAGCTCTGGAATGAATTCACCTGCTCTGAATTGGCCACTGTTGGGGGCGGAGGTG 2190
DB ACTGAGTCTGGAATGAATTCACCTGCTCTGAGTTGGCTTCCTAAATGGGGGCGGAGGTG 1937
QY 2191 TGACTTTGGCTTCCAGGCTTGAAGATTATCTCACCAGCCCTAGCTATATAA-CGGGCTG 2249
DB TTACTTCGGTTCACAGGTTGGAAGATTATCTCACCAGCCCTAGCTATATAAGCTGACCG 1997
QY 2250 GTCTGAGGGGCTCCACAGGCCAGTTCACAGGGTTCATCCACAGAGAGAAAAACATAG 2309
DB GTCTGAGGGGCGCCAGCAGGCCCACTCCAGGGATTCCTTC-CACGACAGAAAAACATAC 2056
QY 2310 A 2310
DB 2057 A 2057

RESULT 4
AAI93508
ID AAI93508 standard; cDNA; 1988 BP.
XX AAI93508;
XX AC AAI93508;
XX 06-NOV-2001 (first entry)
DT Human polynucleotide SEQ ID NO 13568.
DE
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX Homo sapiens.
XX WO200164835-A2.
XX 07-SEP-2001.
XX 26-FEB-2001; 2001WO-US04927.
XX PF
XX 28-FEB-2000; 2000US-0515126.
XX PR
XX 18-MAY-2000; 2000US-0577409.
XX (HYSB-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
XX P-PSDB; AAO13577.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
XX Claim 1; SEQ ID NO 13568; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI93508) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1988 BP; 632 A; 384 C; 472 G; 500 T; 0 other;

Query Match 5.5%; Score 130.2; DB 22; Length 1988;

Best Local Similarity 83.1%; Pred. No. 3.6e-27;

Matches 172; Conservative 0; Mismatches 33; Indels 2; Gaps 2;

QY 2105 TCATCTCCCTCTTGGGCTTCCAGACACTAAGCTCTGGAATGAAAATTACCTGCCTCTGA 2164

Db 27 TCCCTCCCTCTTCCAGCTTCCAGACACTGATTTCTGGAATGAAAATTACCTGCCTCTGA 86

QY 2165 ATTGGCCACTGTGGGGCAGGGGTGTGACTTGGCTTCCAGGCTGGAAGATTATCTCAC 2224

Db 87 GTTGGCTCCTAATGGGGCGGGAGTGTACTTCGGTTCCAGTTGGAAGATTATCTCAC 146

QY 2225 CCAGCCCTAGCTATATAA-CGGGCTGGTGTGGAGGGCTCCACAGGGCCAGTTCCAGGGG 2283

Db 147 CCGGCCCCAGCTATATAAGCTGACCGGTGTGGAGGGGCCAGCAGGGCCAACTCCAGGGA 206

QY 2284 TTCTCCACAGAGAGAAAAACATAGA 2310

Db 207 TTCTTTC-CACGACAGAAAAACATACA 232

RESULT 5

AAH02910

ID AAH02910 standard; DNA; 1901 BP.

XX AAH02910;

DT 15-JUN-2001 (first entry)

DE Human shear stress-response coding sequence SEQ ID NO: 73.

XX Human; shear stress-response protein; vascular disease;

KW arteriosclerosis; ds.

XX Homo sapiens.

OS WO200125427-A1.

PN 12-APR-2001.

PD 02-OCT-2000; 2000WO-JP06840.

XX 01-OCT-1999; 99JP-0280976.

XX (KYOW) KYOWA HAKKO KOGYO KK.

PA (NOJI/) NOJIMA H.

PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;

PI Kuga T, Sekine S, Nakamura Y, Sugano S;

XX WPI; 2001-266308/27.

DR P-PSDB; AAB90787.

XX DNA sequences, proteins encoded by them and antibodies against them

PT useful in diagnosis and treatment of vascular disease caused by

PT arteriosclerosis -

XX Claim 20; Page 422-425; 678pp; Japanese.

XX The present invention provides the protein and coding sequences of a

Best Local Similarity 82.6%; Pred. No. 1e-26;

Matches 171; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

QY 2105 TCATCTCCCTCTTGGGCTTCCAGACACTAAGCTCTGGAATGAAAATTACCTGCCTCTGA 2164

Db 27 TCCCTCCCTCTTCCAGCTTCCAGACACTGATTTCTGGAATGAAAATTACCTGCCTCTGA 86

QY 2165 ATTGGCCACTGTGGGGCAGGGGTGTGACTTGGCTTCCAGGCTGGAAGATTATCTCAC 2224

Db 87 GTTGGCTCCTAATGGGGCGGGAGTGTACTTCGGTTCCAGTTGGAAGATTATCTCAC 146

QY 2225 CCAGCCCTAGCTATATAA-CGGGCTGGTGTGGAGGGCTCCACAGGGCCAGTTCCAGGGG 2283

Db 147 CCGGCCCCAGCTATATAAGCTGACCGGTGTGGAGGGGCCAGCAGGGCCAACTCCAGGGA 206

QY 2284 TTCTCCACAGAGAGAAAAACATAGA 2310

Db 207 TTCTTTC-CACGACAGAAAAACATACA 232

RESULT 6

AAD27217

ID AAD27217 standard; DNA; 1901 BP.

XX AAD27217;

XX 09-APR-2002 (first entry)

DE Human CAA58676 protein encoding EST clone X83703 DNA.

XX Human; congestive heart failure; dilative cardiomyopathy; sudden death;

KW hypertrophic cardiomyopathy; ischaemic cardiomyopathy; rhythm disorder;

KW heart muscle disease; conduction disorder; coronary heart disease;

KW systemic arterial hypertension; pulmonary hypertension; endocarditis;

KW pulmonary heart disease; valvular heart disease; pericardial disease;

KW congenital heart disease; gene therapy; syncope; transgenic animal;

XX expressed sequence tag; EST; clone X83703; CAA58676 protein; ds.

XX Homo sapiens.

OS Key

FT Location/Qualifiers

FT CDS

FT 250..1209

FT /*tag= a

FT /product= "Human CAA58676 protein"

FT misc_feature

FT 198..431

FT /*tag= b

FT /note= "66268 cDNA fragment"

FT misc_feature

FT 1604..1753

FT /*tag= b

FT misc_feature

FT 1302..1306

FT /*tag= b

FT /note= "AU-rich mRNA decay element"

FT misc_feature

FT 1391..1401

FT /*tag= b

FT /note= "AU-rich mRNA decay element"

FT misc_feature

FT 1415..1423

FT /*tag= b

FT /note= "AU-rich mRNA decay element"

XX WO200192567-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-EP06165.

XX 30-MAY-2000; 2000US-207400P.

XX (MEDI-) MEDIGENE AG.

XX Bunk D, Reuner B, Beck J, Henkel T;

XX WPI; 2002-122073/16.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DR P-PSDB; AAE16633.

XX Identifying a subject at risk for a heart disease e.g. congestive heart
 PT failure, dilative cardiomyopathy, heart muscle disease, by quantifying
 PT the polypeptide expressed by genes abnormally expressed in heart tissue
 PT
 XX

PS Claim 2a; Fig 10b; 154pp; English.

XX The patent discloses novel target genes abnormally expressed in heart
 CC tissues and their corresponding proteins. The invention also relates to
 CC methods for assessing the expression level of these genes. The method
 CC is used for testing the predisposition of mammals and preferably humans
 CC for a heart disease or for an acute state of such a disease. It is also
 CC useful to treat diseases of the heart such as congestive heart failure,
 CC dilative cardiomyopathy, hypertrophic cardiomyopathy, ischaemic cardio-
 CC myopathy, specific heart muscle disease, rhythm and conduction disorders,
 CC syncope and sudden death, coronary heart disease, systemic arterial
 CC hypertension, pulmonary hypertension, pulmonary heart disease, valvular
 CC heart disease, congenital heart disease, pericardial disease and
 CC endocarditis. Sequences of the invention are also used in gene therapy.
 CC A transgenic non-human mammal comprising the sequences of the invention
 CC are useful for the development for medicaments for the treatments of
 CC heart diseases. The present DNA sequence is expressed sequence tag
 CC (EST) clone X83703 which encodes CAA58676 protein.

XX Sequence 1901 BP; 592 A; 378 C; 460 G; 471 T; 0 other;

Query Match 5.5%; Score 128.6; DB 24; Length 1901;
 Best Local Similarity 82.6%; Pred. No. 1e-26;
 Matches 171; Conservative 0; Mismatches 34; Indels 2; Gaps 2;
 QY 2105 TCATCTCCCTCTTGGGCTTCCAGACACTAAGTCTGGAATGAAATTCACCTGCCTCTGA 2164
 Db 27 TCCCTCCCTCTTTCAGCTTCCAGACACTGATCTGGAATGAAATTCACCTGCCTCTGA 86
 QY 2165 ATTGGCCACTGTGGGGGAGGGGTGTGACTTGGCTTCCAGGTGGAAGATTATCTCAC 2224
 Db 87 GTTGGCTCTTAATGGGGGTGGGAGTGTACTTCGGTTCCAGGTGGAAGATTATCTCAC 146
 QY 2225 CCAGCCCTAGCTATATAA-CGGGCTGGTGGAGGGGCTCCAGAGGCCAGTTCACAGGGG 2283
 Db 147 CCGGCCCTAGCTATATAGCTGACCGGTGGAGGGGCCCCAGCAGGGCCCACTCCAGGGA 206
 QY 2284 TTCATCCACAAGAGAGAAAAACATAGA 2310
 Db 207 TTCCTTC-CACGACAGAAAAACATACA 232

RESULT 7

ABZ24607
 ID ABZ24607 standard; DNA; 7280 BP.

XX AC ABZ24607;

XX 07-APR-2003 (first entry)

DE Mouse gastric antrum mucosal protein-18 gene.

XX Mouse; gastrophine; AMP-18; gastric antrum mucosal protein; mitogen;
 KW growth factor; vulnerary; antiulcer; gene; ds.

XX Mus sp.

PH Key Location/Qualifiers
 FT promoter 1..1874

FT /tag= a

FT CDS 1945..6542

FT /tag= b

FT /product= "Mouse pre-AMP-18"

FT /note= "contains introns"

FT exon 1906..1956

FT /tag= c

FT intron /number= 1
 FT 1957..3531
 FT /tag= d
 FT /number= 1
 FT 3532..3582
 FT /tag= e
 FT /number= 2
 FT 3583..3672
 FT /tag= f
 FT /number= 2
 FT 3673..3813
 FT /tag= g
 FT /number= 3
 FT 3814..4594
 FT /tag= h
 FT /number= 3
 FT 4595..4705
 FT /tag= i
 FT /number= 4
 FT 4706..5607
 FT /tag= j
 FT /number= 4
 FT 5608..5749
 FT /tag= k
 FT /number= 5
 FT 5750..6444
 FT /tag= l
 FT /number= 5
 FT 6445..6542
 FT /tag= m
 FT /number= 6
 FT 6636..6641
 FT /tag= n
 FT polyA_signal
 FT WQ200292758-A2.
 XX 21-NOV-2002.
 XX 29-MAR-2002; 2002WO-US10148.
 XX 29-MAR-2001; 2001US-0821726.
 XX (UYCH-) UNIV CHICAGO.
 XX Toback FG, Martin TE, Powell CT, Agarwal K;
 WPI; 2003-120666/11.
 XX P-PSDB; ABP58258.

XX Gastric Antrum Mucosal Protein 18, useful for preparing a composition
 PT for healing of the injured gastrointestinal tract
 XX Disclosure; Fig 4; 67pp; English.

XX The present sequence is that of a gene encoding a novel member of
 CC the murine gastrophine group, designated antrum mucosal protein 18
 CC (AMP-18). AMP-18 was initially detected in mammalian gastric
 CC antrum mucosa by a differential screen of cDNA libraries obtained
 CC from different regions of the pig stomach. A cDNA was also
 CC isolated from a mouse library (see ABZ24608). Genomic AMP-18 DNA
 CC sequences were subsequently cloned as a prelude to the analysis of
 CC gene regulatory elements. AMP-18 protein is expressed at high
 CC levels only in the gastric antrum, and is synthesised in the
 CC luminal surface mucosal cells. Partially purified AMP-18s from
 CC mouse and pig antrum tissue are mitogenic to confluent stomach
 CC and kidney epithelial cells in culture. This effect is inhibited
 CC by specific antisera, suggesting that AMP-18, or its products, is a
 CC growth factor. AMP-18 may play a critical role in the repair of
 CC the stomach epithelium following damage by agents such as alcohol,
 CC non-steroidal antiinflammatory drugs, or pathogens, particularly
 CC Helicobacter pylori, a causative agent of gastric ulcers and
 CC possibly cancers. AMP-18 protein, or an active peptide of AMP-18,
 CC is used in a claimed method of stimulating the growth of epithelial

CC mammals containing a DNA expression cassette comprising the neuron-
CC specific promoter are useful as laboratory models for studying the
CC function of the Presenilin gene, and for studying the etiology of
CC Alzheimer's disease.

XX SQ Sequence 48974 BP; 12579 A; 10749 C; 11189 G; 14438 T; 19 other;
Query Match 3.8%; Score 88.6; DB 20; Length 48974;
Best Local Similarity 69.2%; Pred. No. 3.5e-14; Indels 1; Gaps 1;
Matches 135; Conservative 0; Mismatches 59;
QY 191 GCGTCATTAGGGCAGCAGTCTCAACCTGGGGGCTCGACCCCTTTGGGGGAATCAAA 250
DB 15385 GCTACACTAGATCGGTGGTCTCAAGCTGGGGTGGGACCCCTTTGGGGGTTGGACA 15326
QY 251 CGACCTTTACAGGGGTCAATATCATCTATCTATATGTCAGGTATTTACATTTACGATT 310
DB 15325 A-CCTTTTCAACAGGTCACACATCAGAGATCCTGCATATCCGATATTTATGTTATGATT 15267
QY 311 CGTACAGTAGCAAAATTCAGGTATGAAATAGCAATGAAATAATTTATGATTGAAGGT 370
DB 15266 CATAACAGCAACAGAAATACAGTTAGGAAGTAGAAATGAAATAATGTTATGATTGGGCG 15207
QY 371 CACCACACATGAGG 385
DB 15206 CACCATGACATGAAG 15192

RESULT 12
ABT09109
ID ABT09109 standard; DNA; 707 BP.
XX AC ABT09109;
XX DT 05-DEC-2002 (first entry)
XX DE Phase-1 Rat CT gene SEQ ID No 197.
XX KW Rat; toxicity study; rat toxic response gene; toxicological response;
XX KW drug development; phase-1 rat CT gene; ds.
XX OS Rattus sp.
XX PN WO200266682-A2.
XX PD 29-AUG-2002.
XX PF 29-JAN-2002; 2002WO-US02935.
XX PR 29-JAN-2001; 2001US-264933P.
XX PR 26-JUL-2001; 2001US-308161P.
XX (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX PA Farris G, Hicken SH, Farr SB;
XX PI WPI; 2002-674961/72.
XX DR

XX Evaluating the toxicity of an agent, useful in drug development or in
XX determining toxicological responses to a new drug, by determining the
XX expression of rat toxicologically relevant genes in the test animal in
XX response to the test agent -
XX Disclosure; Page 155; 388pp; English.
XX The invention relates to a method used for evaluating the toxicity of an
XX agent comprising determining the expression of a rat toxic response
XX gene(s) in the test animal in response to the agent. The method is useful
XX in drug development, particularly for conducting toxicity studies and
XX analysis before a new drug or compound is approved for human consumption
XX or use. The method is also useful in determining toxicological responses
XX to a new drug. This polynucleotide sequence represents a phase-1 rat CT
XX gene of the invention.

XX SQ Sequence 707 BP; 205 A; 174 C; 156 G; 169 T; 3 other;
Query Match 3.5%; Score 83.6; DB 24; Length 707;
Best Local Similarity 67.3%; Pred. No. 9.9e-14;
Matches 148; Conservative 0; Mismatches 69; Indels 3; Gaps 2;
QY 171 GAGCTTCTCTTCAGGATCAGCCTGATTTCTAGGGCAGCAGTTCTCAACCTGGGGGCTCG 230
DB 232 GAGCCATCCCTCCAGCCCGAGCTGTTTTATGGAAGTGATTTCTCAACTCATGGTTCATG 291
QY 231 ACCCTTTTGGGGG--AATCAAAACGACCCCTTTTACAGGGGTCAATATCTATCTATATAT 288
DB 292 ACCCTTTTGGGGGTTAAATGACCCCTTTTACATATCAAAATCAATCAATACCCCTGCAG 351
QY 289 GTGAGGTATTTACATTAGATTGCTGTAACAGTAGAGCAAAATTACAGGTATGAAATAGCAATG 348
DB 352 AGCAGATATTTCATTTGCAATTCGAATCCGCAACAGCAGCAAAATTACAGTTACGAGTAGCAAAAG 411
QY 349 -AAATAATTTTATGATTGAAGGTCAACCAACATGAGGCC 387
DB 412 AAATAATTTTACGGTTGAAGGTCAACACACGAGGAC 451

RESULT 13
ABX16390/c
ID ABX16390 standard; DNA; 659158 BP.
XX AC ABX16390;
XX DT 09-APR-2003 (first entry)
XX DE Mouse high growth region.
XX KW High growth region; high growth phenotype; Socs2; body size;
XX KW suppressor of cytokine signaling 2; ds; mouse.
XX OS Mus sp.
XX PN US2002155564-A1.
XX PD 24-OCT-2002.
XX PF 26-JAN-2001; 2001US-0771208.
XX PR 29-DEC-1997; 97US-0999477.
XX (REGC) UNIV CALIFORNIA.
XX Medrano JF, Bradford E, Horvat S;
XX WPI; 2003-182637/18.

XX Novel gene that when downregulated or knocked-out, results in high
XX growth phenotype, useful for regulating body size in mammals e.g.
XX rodent, bovine and canine -
XX Disclosure; SEQ ID NO 18; 49pp; English.
XX The invention describes an isolated nucleic acid molecule encoding a
XX gene product that, when knocked out, results in a high growth (hg)
XX phenotype. For example a nucleic acid disrupting the Socs2 gene is
XX useful for producing an animal characterised by a hg phenotype, by
XX inhibiting expression of Socs2 (suppressor of cytokine signaling 2)
XX gene. The nucleic acids of the invention are useful for regulating
XX body size in mammals. gene. The nucleic acids of the invention
XX are useful for regulating body size in mammals. This sequence
XX represents the mouse high growth region.
XX Note: This sequence did not form part of the printed specification
XX but was obtained in electronic format directly from the US patent
XX office at
XX seqdata.uspto.gov/sequence.html?DocID=20020155564.

CC	useful for monitoring the efficacy of a drug against depression.
XX	
SQ	Sequence 10917 BP; 2423 A; 2842 C; 2929 G; 2706 T; 17 other;
Query Match	3.4%; Score 80.4; DB 24; Length 10917;
Best-Local Similarity	70.8%; Pred. No. 3.9e-12;
Matches 136; Conservative	0; Mismatches 51; Indels 5; Gaps 2;
QY	197 TTTACGGGACGAGTCTTCAACCTCTGGGGCCTCGACCCCTTTGGGGGAATCAAAAGCACCC 256
Db	10340 TGGGAGGACAGAGGTTCTCAACCTGTGGGGACAAACCCCTTTGGGG--TTTGAATGACTT 10397
QY	257 TTTACAGGGGTCACATATCA---TCTATCCTTATATGTCAGGTATTTACATTACGATTGGT 313
Db	10398 TTTACAGGGGTTGCTTTGCCAAGACCATCAGAAACACAGGTATTTACATTCCCAATTGAT 10457
QY	314 AACAGTAGCAAAATTTACAGGTATGAAATAGCAATGAAATAATTTTATGATTGAAGGTGAC 373
Db	10458 GACAGTAGTAAATTTACAGTTATGAGTAGCAATGAAATAATTTTATGTTGGGGTGAC 10517
QY	374 CACAACATGAGG 385
Db	10518 CACATCATGAAG 10529

RESULT 15	
AAV06155	
ID	AAV06155 standard; DNA; 892 BP.
XX	
AC	AAV06155;
XX	
DT	06-MAY-1998 (first entry)
XX	
DE	Viral infection gene SEQ ID NO:75.
XX	
XX	
KW	Viral infection; tumour suppressor; cellular gene; rat; cancer;
KW	serum protein; inhibitor; malignant phenotype; HIV; influenza;
KW	hepatitis; retrovirus; immunodeficiency; ds.
XX	
OS	Rattus sp.
XX	
PN	W09739119-A1.
XX	
PD	23-OCT-1997.
XX	
PF	11-APR-1997; 97WO-US06067.
XX	
PR	15-APR-1996; 96US-0015334.
XX	
PA	(UYVA-) UNIV VANDERBILT.
XX	
PI	Dubois RN, Organ EL, Rubin DH;
XX	
DR	WPI; 1997-526456/48.
XX	
PT	Genes involved in viral infection and tumour suppression - used to
PT	develop products for reducing or preventing viral infection or for
PT	suppressing tumours
XX	
PS	Claim 1; Page 82-83; 101pp; English.

The present sequence represents a viral infection gene. The present invention describes nucleic acid sequences isolated from rat. The sequences of the invention comprise 70 viral infection (VI) genes and 8 tumour suppressor (TS) genes. Propagating cell cultures in the absence of the serum protein (SP) allows selective elimination of cells persistently infected with a virus from the cell culture. Inhibitors of the SP can be used for reducing or inhibiting a viral infection. Inhibitors of the TS gene products can be used to suppress a malignant phenotype (MP). The methods and inhibitors can be used with viruses such as HIV, influenza, hepatitis virus or animal retroviruses such as simian immunodeficiency virus, avian immunodeficiency virus, bovine immunodeficiency virus, feline immunodeficiency virus, equine infectious

CC anemia virus, caprine arthritis encephalitis virus or visna virus.
CC Because the identified genes are non-essential to cell survival, the
CC treatment methods can be used in subjects without serious detrimental
CC effects to the subjects.

XX

SQ Sequence 892 BP; 224 A; 206 C; 185 G; 229 T; 48 other;

Query Match 3.3%; Score 78; DB 18; Length 892;
Best Local Similarity 70.5%; Pred. No. 5e-12; Mismatches 55; Indels 7; Gaps 3;
Matches 148; Conservative 0;
QY 177 CTCCTTCAGGATCAGCCTGATTCTAGGGCAGCAGTTCTCAACCTGGGGCCTCGACCCCT 236
DB 119 CACGTACATAGATAGTCAAAATCTAGAGCACTGTTCTATACCTGTGAGTTGCAACCCCT 178
QY 237 TTGGG---GGAATCAAACGACCCT-TTACAGGGGTCACATATCATCTATCTATATGTCA 292
DB 179 TTGGGAGTGGGTCAAATGACCCTATCACAGGGGTCTCAAATGAGATATCTCTGCATATCA 238
QY 293 GGTATTATACATTACGATTCGTAACAGTAGCAAAATACAGGTATGAAATAGCAATCAAT 352
DB 239 AATATTATCATATGATTTCATAGTAGTACCAGAAATACAGTTATGAAGTTACA---AAAT 295
QY 353 AATTTTATGATTGAAGGTCACCAACACATG 382
DB 296 AATTTTATAGCTGAGAGTCACCACACATG 325

Search completed: November 14, 2003, 04:10:52
Job time : 613.653 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 03:46:27 ; Search time 4902.75 Seconds
(without alignments)
11689.358 Million cell updates/sec

Title: US-10-005-337A-1
Perfect score: 2358
Sequence: 1 ggaaccccttcattgttaaca.....caggtcgaggccaccatgg 2358

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_esti.*
9: gb_esti.*
10: gb_est2.*
11: gb_est3.*
12: gb_est4.*
13: gb_est5.*
14: gb_est6.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pin.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gssi.*
29: gb_gss2.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	585	24.8	600	28	BH043261
2	505.4	21.4	586	28	AZ295534
3	306.8	13.0	371	13	BY091041
4	269.6	11.4	353	13	BY079997

C	5	191.6	8.1	478	12	BI293043
C 6	183	7.8	434	10	BF543101	
C 7	171	7.3	488	9	AW252440	
C 8	171	7.3	511	9	AW251306	
C 9	167.8	7.1	370	9	AI710529	
C 10	131.8	5.6	735	9	AUI139209	
C 11	119.6	5.1	509	10	BE627514	
C 12	116.2	4.9	538	28	AZ804735	
C 13	116	4.9	522	28	AQ480395	
C 14	114.6	4.9	518	28	AZ114168	
C 15	113	4.8	596	28	AZ290439	
C 16	112.4	4.8	623	10	BB470798	
C 17	108	4.6	577	28	AZ734079	
C 18	107.8	4.6	556	28	AZ373614	
C 19	107.8	4.6	589	28	BH360537	
C 20	107.2	4.5	839	28	BZ110003	
C 21	107	4.5	369	28	AZ600690	
C 22	106.2	4.5	4565	11	AK052332	
C 23	106	4.5	3534	11	AK038950	
C 24	104	4.4	310	28	AZ939809	
C 25	103.6	4.4	585	14	CA882248	
C 26	103.6	4.4	834	28	BZ185566	
C 27	103.2	4.4	519	28	AZ488841	
C 28	103.2	4.4	640	29	BZ239325	
C 29	103	4.4	574	28	AZ563281	
C 30	102.6	4.4	422	10	BF412836	
C 31	102.6	4.4	810	29	BZ264838	
C 32	102	4.3	364	28	AZ348698	
C 33	101.8	4.3	774	28	BZ189630	
C 34	101.6	4.3	676	28	AZ496223	
C 35	101.2	4.3	594	28	AZ389177	
C 36	101	4.3	477	28	AZ778894	
C 37	101	4.3	596	28	AZ856151	
C 38	100.8	4.3	407	13	BY549634	
C 39	100.6	4.3	600	12	BG803092	
C 40	100.4	4.3	591	28	AZ445982	
C 41	100.4	4.3	787	28	BZ128535	
C 42	100.4	4.3	803	29	BZ234545	
C 43	100.4	4.3	2591	11	AK036783	
C 44	100.2	4.2	1085	29	BZ223957	
C 45	100	4.2	473	28	AZ060257	

ALIGNMENTS

RESULT 1
BH043261
LOCUS
DEFINITION
RPCI-24-330A6.TJ RPCI-24 Mus musculus genomic clone RPCI-24-330A6,
genomic survey sequence.

ACCESSION
BH043261
VERSION
BH043261.1
KEYWORDS
GSS.

SOURCE
Mus musculus (house mouse)

ORGANISM

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 600)

TITLE
JOURNAL
COMMENT
Mouse BAC End Sequences from Library RPCI-24
Unpublished
Other_GSSs: RPCI-24-330A6.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@igr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC

BI293043 UI-R-DK0-
BF543101 UI-R-AGL-
AW252440 UI-R-BJ0-
AW251306 UI-R-BJ0-
AI710529 UI-R-AGL-
AUI139209 AUI139209
BE627514 u052608.Y
AZ804735 2M006511.Y
AQ480395 RPCI-11-2
AZ114168 RPCI-23-4
AZ290439 RPCI-23-5
BB470798 BB470798
AZ734079 RPCI-24-1
AZ373614 1M0125N21
BH360537 CH230-164
BZ110003 CH230-164
AZ600690 1M0418D12
AK052332 Mus muscu
AK038950 Mus muscu
AZ939809 2M0198P19
CA882248 K0996807-
BZ185566 CH230-340
AZ488841 1M0319O13
BZ239325 CH230-282
AZ563281 RPCI-23-2
BF412836 UI-R-BT1-
BZ264838 CH230-315
AZ348698 1M0085D24
BZ189630 CH230-525
AZ496223 1M0332H21
AZ389177 1M0149M20
AZ778894 2M0014H04
BY549634 2M0160A02
BG803092 0194-80 M
AZ445982 1M0242B09
BZ128535 CH230-298
BZ234545 CH230-427
AK036783 Mus muscu
BZ223957 CH230-446
AZ060257 RPCI-23-4

library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.html>). BAC end page: http://www.tigr.org/tadb/bac_ends/mouse/bac_end_intro.html
Plate: 330 row: A column: 6
Seq primer: SP6
Class: BAC ends.

FEATURES

source
Location/Qualifiers
1..600
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-330A6"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

BASE COUNT 158 a 139 c 142 g 161 t

Query Match 24.8%; Score 585; DB 28; Length 600;

Best Local Similarity 100.0%; Pred. No. 1.4e-148; Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 TTCTCAACTGGGGGCTCGACCCCTTTGGGGGAATCAACGACCCCTTTACAGGGGTAC 270
DB 16 TTCTCAACTGGGGGCTCGACCCCTTTGGGGGAATCAACGACCCCTTTACAGGGGTAC 75
QY 271 ATATCATCTATCTATATCTAGGTATTTACATTACGATTCGTAACAGTAGCAAAATTAC 330
DB 76 ATATCATCTATCTATCTAGGTATTTACATTACGATTCGTAACAGTAGCAAAATTAC 135
QY 331 AGGTATGAATAGCAATGAATTAATTTATGATTGAGGTACCACCAATGAGGCCGCC 390
DB 136 AGGTATGAATAGCAATGAATTAATTTATGATTGAGGTACCACCAATGAGGCCGCC 195
QY 391 ACACCTGTTCTAGAGAAAATACCTGGTGGGAAAGTTTGGGAAAGCTTTCTGTCCA 450
DB 196 ACACCTGTTCTAGAGAAAATACCTGGTGGGAAAGTTTGGGAAAGCTTTCTGTCCA 255
QY 451 TTCTTCATTCTTCAAGTGTATGTTTACAGAAAGCTTTACAGTCTTTCTGTGGGCTC 510
DB 256 TTCTTCATTCTTCAAGTGTATGTTTACAGAAAGCTTTACAGTCTTTCTGTGGGCTC 315
QY 511 TTAGTAAGTCTGAGTAGGAACTGTATGTACCAGGTCTGCTTCTTATGGGTGGAGCAAGA 570
DB 316 TTAGTAAGTCTGAGTAGGAACTGTATGTACCAGGTCTGCTTCTTATGGGTGGAGCAAGA 375
QY 571 CGCATCGTGGGTGGAGCAAGCAACCTCTAGCTCTGATCTGATCATCATAGCAAGTA 630
DB 376 CGCATCGTGGGTGGAGCAAGCAACCTCTAGCTCTGATCTGATCATCATAGCAAGTA 435
QY 631 GCTTAATGTTTCTGTCTAGGTGTCTCTGTGTGAATCGAGATCTTGGCCCTTGTGGA 690
DB 436 GCCTAATGTTTCTGTCTAGGTGTCTCTGTGTGAATCGAGATCTTGGCCCTTGTGGA 495
QY 691 ATTAGGAGGACCAAAATCTAGAGATTCAGACTGCTCAGCAGCCAGTCTCTTCTCT 750
DB 496 ATTAGGAGGACCAAAATCTAGAGATTCAGACTGCTCAGCAGCCAGTCTCTTCTCT 555
QY 751 CAAAGGAAAGGTCTCAACTCTCAGCCCCCTTAGCTCTGAGTCAG 795
DB 556 CAAAGGAAAGGTCTCAACTCTCAGCCCCCTTAGCTCTGAGTCAG 600

RESULT 2
AZ295534
LOCUS 586 bp DNA linear GSS 27-JUL-2000

DEFINITION

RPCI-23-105P3.TJ RPCI-23 Mus musculus genomic clone RPCI-23-105P3, genomic survey sequence.

ACCESSION AZ295534.1 GI:9537319

VERSION GSS.

KEYWORDS Mus musculus (house mouse)

SOURCE

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 586)
Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret ,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-23

Unpublished

Other_GSSs: RPCI-23-105P3.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)

or from Resea ch Genetics (info@resgen.com). BAC end page:

http://www.tigr.org/tadb/bac_ends/mouse/bac_end_intro.html

Plate: 105 row: P column: 3

Seq primer: SP6

Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..586

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-105P3"

/sex="Female"

/lab_host="DH10B"

/clone_lib="RPCI-23"

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site 1:

ECORI; Site 2: EORI; Female C57BL/6J mouse kidney and/or

brain genomic DNA was isolated and partially digested

with a combination of EORI and EORI Methyase. Size

selected DNA was cloned into the pBACE3.6 vector at the

ECORI sites. The ligation products were transformed into

DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 141 a 148 c 154 g 143 t

ORIGIN

Query Match 21.4%; Score 505.4; DB 28; Length 586;

Best Local Similarity 98.5%; Pred. No. 8.6e-127;

Matches 531; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 1765 GTGGCACCAGACCCATCTCCCTCGGGCTGATATCCCAAGATAGGATGCTCC 1824

DB 1 GTGGCACCAGACCCATCTCTTCCCTCGGGCTGATATCCCAAGATAGGATGCTCC 59

QY 1825 AAAGCAACACTTCCAGCCAACTGGAGTCTGATAGTCCAGTTATCAGAAAGATATGCG 1884

DB 60 AAAGCAACACTTCCAGCCAACTGGAGTCTGATAGTCCAGTTATCAGAAAGATATGCG 119

QY 1885 TGTAAGTGTGATGCACAGTGTTCATTTTCTTGATACGTTAGTCAATATGAGAGCTGACA 1944

DB 120 TGTAAGTGTGATGCACAGTGTTCATTTTCTTGATACGTTAGTCAATATGAGAGCTGACA 178

QY 1945 AAGAGGAAAAAGAGCAGCGATGTGGTGAATATTAACAGGAGCTGTCCTCCCTGGCTTC 2004

DB 179 AAGAAGGAAAAAGAGCAGCGATGTGGTGAATATTAACAGGAGCTGTCCTCCCTGGCTTC 238

QY 2005 CGATACGTGGGATGACTCGCATTTGCTGAGCGGTGTGGTCACTGCCAAAGGAATGACCTC 2064

Db 239 CGATACGGGATGACTGCGATTGCTAGCGGTGTGGTCACTGCCAAGGAATACACCTC 298
Qy 2065 TCACATTTCTTCTGATTCGCATACGCGCGCGCCAGCTTGTGCATCTCCCTCTTGGGCTTC 2124
Db 299 TCACATTTCTTCTGATTCGCATACGCGCGCGCCAGCTTGTGCATCTCCCTCTTGGGCTTC 358
Qy 2125 CCAGACACTAAGTCTGGAATGAATAATCACTGCTCTGTAATGGCCACTGGTGGGGCA 2184
Db 359 CCAGACACTAAGTCTGGAATGAATAATCACTGCTCTGTAATGGCCACTGGTGGGGCA 418
Qy 2185 GGGGTGTGACTTGGCTTCCAGGCTGGAAGATTATCTCACCCGCCCTAGCTATATAACG 2244
Db 419 GGGGTGTGACTTGGCTTCCAGGCTGGAAGATTATCTCACCCGCCCTAGCTATATAACG 478
Qy 2245 GGCTGGTGTGGAGGGGCTCCACAGGCGCAGTTCACAGGGTTCATCCACAGAGAGAAAA 2303
Db 479 GGCTGGTGTGGAGGGGCTCCACAGGCGCAGTTCACAGGGTTCATCCACAGAGAGAAAA 537

RESULT 3

BY091041
LOCUS BY091041 371 bp mRNA linear EST 07-DEC-2002
DEFINITION BY091041 RIKEN full-length enriched, 10 days neonate heart Mus
musculus cDNA clone K630084E21 5', mRNA sequence.

ACCESSION

BY091041

VERSION

EST.

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

1 (bases 1 to 371)

Okazaki, Y., Furuno, M., Kasakawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.

Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,

Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,

Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani,

L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,

A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,

Gough, G., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,

Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierzki, R.M.,

King, B.L., Kongaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,

P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,

H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G.,

Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,

Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring,

B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempie, C.A., Setou,

M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale,

R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,

Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa,

M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,

Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,

M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,

Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,

Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,

K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,

E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12468851

CONTACT: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gs.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane,
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A.,
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.: Direct
Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES

source
1..371
Location/Qualifiers
/organism="Mus musculus"
/mol_type="cDNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K630084E21"
/tissue_type="heart"
/dev_stage="10 days neonate"
/clone_lib="RIKEN full-length enriched, 10 days neonate
heart"
BASE COUNT 91 a 94 c 105 g 81 t
ORIGIN

Query Match 13.0%; Score 306.8; DB 13; Length 371;
Best Local Similarity 99.4%; Pred. No. 1.8e-72;
Matches 308; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2008 TAGCTGGGATGACTGCGATTGCTGAGCGGTGTGGTCACTGCCAAGGAATGACCTCTCA 2067
Db 1 TAGCTGGGATGACTGCGATTGCTGAGCGGTGTGGTCACTGCCAAGGAATGACCTCTCA 60

Qy 2068 CATTTCTTCTGATTGCGATACGCCGGGCCAGCTTGTCTCTCTTTGGGCTTCCCA 2127
Db 61 CATTTCTTCTGATTGCGATACGCCGGGCCAGCTTGTCTCTCTTTGGGCTTCCCA 120

Qy 2128 GACACTAAGTCTGGAATGAAATTCACCTGCTCTGAATTTGGCCACTGTTGGGGGCGAGG 2187
Db 121 GACACTAAGTCTGGAATGAAATTCACCTGCTCTGAATTTGGCCACTGTTGGGGGCGAGG 180

Qy 2188 GTGTGACTTGGCTTCCAGGCTGGAAGATTATCTCACCGAGCTAGCTATATAACGGGC 2247
Db 181 GTGTGACTTGGCTTCCAGGCTGGAAGATTATCTCACCGAGCTAGCTATATAACGGGC 240

Qy 2248 TGGTGTGGAGGGGCTCCACAGGGCCAGTTCAGGGGTTTCATCCACAGAGAGAAAAACAT 2307
Db 241 TGGTGTGGAGGGGCTCCACAGGGCCAGTTCAGGGGTTTCATCCACAGAGAGAAAAACAT 300

Qy 2308 AGACTCGAGG 2317
Db 301 AGACTCAGG 310

RESULT 4

BY079997 353 bp mRNA linear EST 06-DEC-2002
LOCUS BY079997 RIKEN full-length enriched, 10 days neonate heart Mus
DEFINITION musculus cDNA clone K630022E19 5', mRNA sequence.
ACCESSION BY079997

```

VERSION BY079997.1 GI:26185447
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
AUTHORS Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest
A., Frazer, K. S., Gasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Gimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,
King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons
P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki
H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perle, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C. A., Setou
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale
R. D., Tomita, M., Vezardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A.,
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami
M., Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y. Direct
Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
Location/Qualifiers
1. .353

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K630022E19"
/tissue_type="heart"
/dev_stage="10 days neonate"
/clone_lib="RIKEN full-length enriched, 10 days neonate
heart"
BASE COUNT 82 a 88 c 102 g 79 t 2 others
ORIGIN
Query Match 11.4%; Score 269.6; DB 13; Length 353;
Best Local Similarity 95.5%; Pred. No. 2.8e-62;
Matches 298; Conservative 0; Mismatches 11; Indels 3; Gaps 2;
QY 2066 GATACGTGGGATGACTCGCATTCGTCAGGGGTGGTGGTCTCTCCCAAGGAATGACCCCTCT 2065
Db 2 GATACGTGGGATGACTCGCATTCGTCAGGGGTGGTGGTCTCTCCCAAGGAATGACCCCTCT 61
QY 2066 CACATTTCTCTCTGATTGTCATACCGCGCGGCAGCTTGTCATCTCCCTCTTGGGCTTCC 2125
Db 62 CACATTTCTCTCTGATTGTCATACCGCGCGGCAGCTTGTCATCTCCCTCTTGGGCTTCC 121
QY 2126 CAGACACATAAGTCTGGAATGAAATTCACCTGCTCTGAAATGGCCACTGGTGGGGCAG 2185
Db 122 CAGACACATAAGTCTGGAATGAAATTCACCTGCTCTGAAATGGCCACTGGTGGGGCAG 181
QY 2186 GGGTGTGACTTGGCTTCCAGCTCGAAGATTATCTCACCAGCCCTAGCTATATAACCG 2245
Db 182 GGGTGTGACTTGGCTTCCAGCTCGAAGATTATCTCACCAGCCCTAGCTATATAACCG 239
QY 2246 GCTGTGTGGAGGGCTCCACAGGGCCAGTTCCAGGGTTCATCCACAGAGAGAAAAAC 2305
Db 240 GCTGTGTGGAGGGCTCCACAGGGCCAGTTCCAGGGTTCATCCACAGAGAGAAAAAC 298
QY 2306 ATAGACTCGAGG 2317
Db 299 ATAGACTCAGG 310
RESULT 5
Bi293043/c
LOCUS
DEFINITION
UI-R-DK0-cdj-f-07-0-UI-s1 UI-R-DK0 Rattus norvegicus cDNA clone
UI-R-DK0-cdj-f-07-0-UI-3', mRNA sequence.
BI293043
BI293043.1 GI:14954159
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 478)
AUTHORS
Bonaldo, M. F., Lennon, G. and Soares, M. B.
TITLE
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
97044477
PUBMED
8889548
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the

```


normalized rat heart pool library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cDNA sequence: 170-292, >URR1A#DNA/MER1 type
Seq primer: M13 Forward
POLYA=Yes.

FEATURES

Location/Qualifiers

```

1. .478
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DKO-cd1-f-07-0-UI"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-DKO"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; The UI-R-DKO library is a subtracted library derived from a mixture of five individually tagged normalized rat libraries: brain-nRBP (20%), heart-nRBP (20%), kidney-nRBP (20%), aorta-nRAP (20%), and placenta-nRBP (20%). Each original library was constructed from a mixture of equal amounts of RNA from seven different developmental time-points: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, and adult day 200. (Exception: the aorta pool does not contain embryonic day 17 RNA and the placenta pool contains only the three embryonic stages). Each library was normalized individually according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). For construction of the DKO subtracted library, plasmid DNA from each of the five individually tagged normalized libraries was mixed in the proportions specified above and electroporated into competent bacteria for production of single-stranded circular DNA representing the pool of libraries. Single-stranded circular DNA representing these five normalized libraries was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template preparation) comprising: a) a set of about 1,000 arrayed clones from each of the five non-normalized libraries of brain (CTOs), heart (CSOs), kidney (CUDs), aorta (CWOs), and placenta (CXOs). The resulting pool of approximately 5,000 clones represented about 33.3% of the final driver population. A set of about 2,000 arrayed clones from each of the five normalized libraries of brain (CTO), heart (CSO), kidney (CUO), aorta (CWO), and placenta (CXO). The resulting pool of about 10,000 clones represented about 66.6% of the final driver population.
TAG LIB=UI-R-DKO
TAG TISSUE=rat heart pool
TAG_SEQ=ATGATGATAC"
BASE COUNT 118 a 103 c 110 g 146 t 1 others
ORIGIN
Query Match 8.1%; Score 191.6; DB 12; Length 478;
Best Local Similarity 69.5%; Pred. No. 7.4e-41;
Matches 372; Conservative 0; Mismatches 80; Indels 83; Gaps 5;
QY 1 GGATCCTTTTCATGTTTAACAATATCAACCTTAACCAAGGGGACAGCCTGCCTGACAGT 60
DB 475 GGATCCTCTAATGTTTAACAATATGACAAGCTAGCCCA--GGGGAAGAGCTGCCTGACAGT 418
QY 61 GCGTTTGCCACCCATGAATCTCTAGTCTAGTCGTTGTGAAACTAGCCCATCCCA 120
DB 417 GCGTNTGCCACCCATGAACACATCTAGTCCCTCCCTTGTGAAATTCAGGCCATCCCA 358
QY 121 ACATCTTCTGCAAGCCCATCTCTACAGGTGCTCATTTGGGAATTTTCCTGGAGCTCTCT 180
DB 357 ACA-----GGTCTGTCCTCCGAGAGCTCTCT 333

```

```

QY 181 TTCAGGATCAGCCTGATTCTTAGGGCAGCAGTTCTCAACCTGGGGGCGCTCGACCCCTTTGG 240
DB 332 TTCAGGTTGAGCCTGATTCTTAGGGCAGCAGTTCTCAACCTGGA----- 290
QY 241 GGGAAATCAACGACCCCTTTACAGGGGTACATATCATCTATCTATATATGTCAGGTATTTA 300
DB 289 -----GTCACATATCATCTATCTCTGTATATCAAAATATTTCA 255
QY 301 CATTACGATTTCGTAAACAGTAGCAAAATTACAGGTATGAAATAGCAATGAAATATTTTAT 360
DB 254 CATCATGACTCATAACAGTAACACAAATTACAGGCATGATGTAGCAATGAAATGATTTAT 195
QY 361 GATTGAAGTCCACCAACATGAGGCCGCCACACTGTTCTAGAGAAATAATCACTGGGTG 420
DB 194 GGTGAGGGTCCACCAACATGAGGCCGCCACACTGTTCTGGGGGAA---CAGTGGAGAG 138
QY 421 GGGAAAGGTTTCGGGAAGCCTTTCTGTCATCTCTTCAATCTTCAAAAGTGTGTTTCACA 480
DB 137 AGGAAGGATTGAGGAAGCC-TTCCATCTGTTCTGCACCTCTTCAAAGGGATGATTTCACA 79
QY 481 GAAAGCCTTTTCAGCTGTTCTGCTGGGGCTCTTAGTAAGTCTGAGTAGGAACTGTA 535
DB 78 GAAAGCCTTTGTCGGTCTTTGTTGGGGCTCTTAGTAATCTGAGCAGAACTGCA 24

RESULT 6
BF543101 434 bp mRNA linear EST 11-DEC-2000
LOCUS UI-R-AG1-aal-b-08-0-UI.r1 UI-R-AG1 Rattus norvegicus cDNA clone
DEFINITION UI-R-AG1-aal-b-08-0-UI 5', mRNA sequence.
ACCESSION BF543101
VERSION BF543101.1 GI:11634214
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 434)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
CDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) This clone is also available through the I.M.A.G.E. Consortium at LLNL (info@image.llnl.gov). IMAGE ID= 1790595 The following repetitive elements were found in this cDNA sequence: 45-167, >URR1A#DNA/MER1 type
Seq primer: M13 Forward
Location/Qualifiers
1. .434
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-AG1-aal-b-08-0-UI"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-AG1"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; The UI-R-AG1 library is a normalized library constructed from 13 dpc rat ventricle. The tag is a string of 6 nucleotides

```

present between the Not 1 site and the oligo-dT track.
The library was constructed as described by Bonaldo,
Lennon and Soares, Genome Research 6: 791-806, 1996.
Tissue provided by Jim Lin, Department of Biology,
University of Iowa."

```

BASE COUNT      119 a      94 c      103 g      117 t      1 others
ORIGIN

Query Match      7.8%; Score 183; DB 10; Length 434;
Best Local Similarity 73.0%; Pred. No. 1.6e-38;
Matches 276; Conservative 0; Mismatches 96; Indels 6; Gaps 3;

QY 261 CAGGGGTACATATCATCTATCTATGTCTAGGTATTTACATTAGCATTCGTAAACAGTA 320
DB |||||
QY 43 CTGGAGTCATATCATCTATCTCTGTATATCAAAATATTCACATCATGCTCTATACAGTA 102
DB |||||
QY 321 GCAAAATTACAGTATGAATAGCAATGAATAATTTTATGATTAAGTTCACCAACCA 380
DB |||||
QY 103 ACACAAATTACAGCATGATGTAGCAATGAATGATTTTATGTTGAGGGTCAACCAACA 162
DB |||||
QY 381 TGAGCGCCGACACTGTTCTAGAGAAATATCCTGGGTGGGAAAGGTTTGGGAAGCC 440
DB |||||
QY 163 TGAGGCCACACACTGTTCTGGGGAA---CAGTGGAGAGAGAAAGGATTGAGGAAGCC 219
DB |||||
QY 441 TTCTCTCATCTCTTCATTTCTTCAAGTGTATGTTTCACAGAAAGCCTTTCAGCTGTTCT 500
DB |||||
QY 220 -TTCCATCTGTTCTGACCTCTTCAAGGGATGATTTACAGAAAGCCTTTCGCTGTTT 278
DB |||||
QY 501 GCTGGGCTCTTAGTAAAGTCTAGTAGGAACCTGTATGTACCAAGGCTGCTTCTTATGGGT 560
DB |||||
QY 279 GTTGGGCTCTTAGTAAATCTGAGCAGAAACTGCACCAGGAATTTT--ATCTGGTTTGG 336
DB |||||
QY 561 GGAGCCAAAGCATGCTGGTGGGAGGAGGAGCAACCTCACCCTCTAGCTCTGCATCC 620
DB |||||
QY 337 TCATTTTCTGCTTCTCTCTGTTGGAACCAAGACTGGACTCCATTTGTCTAGCCCTGCATCT 396
DB |||||
QY 621 ATAGCAAGTAGGCTAATG 638
DB |||||
QY 397 GCAGCTAGTTCCTAATG 414
DB |||||

```

```

RESULT 7
AW252440/c
LOCUS
DEFINITION
UI-R-BJ0-ax-e-05-0-UI.s1 UI-R-BJ0 Rattus norvegicus cDNA clone
UI-R-BJ0-ax-e-05-0-UI 3', mRNA sequence.
ACCESSION
AW252440
VERSION
AW252440.1 GI:6596031
KEYWORDS
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 488)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the

```

normalized AV canal at 15 dpc library cDNA Library Preparation:
M.B. Soares Lab Clone distribution: clones will be available
through Research Genetics (www.regen.com) The following repetitive
elements were found in this cDNA sequence: 1-44,
>POLY_A#Simple repeat
Seq primer: M13 Forward
POLYA=Yes.

FEATURES

```

Location/Qualifiers
1..488
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJ0-ax-e-05-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-BJ0"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ0
library is a subtracted library derived from the UI-R-AA1,
UI-R-AB1, UI-R-AC1, UI-R-AD1, UI-R-AE1, UI-R-AF1, and
UI-R-AG1 libraries. These libraries represent tissues from
rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal
at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV
canal at 15 dpc, and ventricle at 13 dpc. The tag is a
string of 5-6 nucleotides present between the Not I site
and the oligo-dT track. The library was constructed as
described by Bonaldo, Lennon and Soares, Genome Research
6: 791-806, 1996.
TAG LIB=UI-R-BJ0
TAG TISSUE=AV canal at 15 dpc
TAG_SEQ=GAAGG"

```

```

BASE COUNT      115 a      111 c      106 g      155 t      1 others
ORIGIN

Query Match      7.3%; Score 171; DB 9; Length 488;
Best Local Similarity 78.4%; Pred. No. 3.3e-35;
Matches 261; Conservative 0; Mismatches 55; Indels 17; Gaps 4;

```

```

QY 879 ATGAAACGAGAAACCAACAGTTATCCATTGATAGCTCTCAGGACAGATAGGACAGAG 938
DB |||||
QY 349 ATGGAACCCAGAAAGCCCAACAGTTGCTCTTCGATAGTGTCTCAGGACACGACAGAG 290
DB |||||
QY 939 AGAACACTAGGAGAGGGAACCCAGAGCAAGGTATTAGTGTCTGTTTTCAGGGC 998
DB |||||
QY 289 ----CACTAGGAGAGAGAACCCAGAGGA-----TATCAGTGTCTGTTTTCAGGGC 239
DB |||||
QY 999 AATGTCTTGTACTGAAGATTCTAGAAACACAAATTTGCTGGTTGAACAGCTGAAGTGGGT 1058
DB |||||
QY 238 AATGCTCATTCGGAAGGTTCTAGAAACACACAGTTTACTGGATGAAGAGCTGAAGTGGGT 179
DB |||||
QY 1059 GGGG---GTTCTTACCCCATGTTTCATGGAAGGTTAGTGAGGAGAGACAGATATATGATG 1115
DB |||||
QY 178 GGGGTAAGAATAAACACCAACCGTTCTCACTGAAGGTTGAGGAGGACGAGACATACGATG 119
DB |||||
QY 1116 GCACGATAACAAACATACACAAACCCCTAATTAACATTCCTCTTCTTACTGACACCCC 1175
DB |||||
QY 118 GCAGCAT-----ACAAACATAACACCCTTAATATGCTTCTCTCTCTGCTACTGACTCT 64
DB |||||
QY 1176 CTTCACTCTCTCTTTTCATAAAAAATAAAAAA 1208
DB |||||
QY 63 CCCTTCACCTCTCTTTCATAAAAAATAAAAAACA 31
DB |||||

```

```

RESULT 8
AW251306/c
LOCUS
DEFINITION
UI-R-BJ0-adg-f-04-0-UI.s1 UI-R-BJ0 Rattus norvegicus cDNA clone
UI-R-BJ0-adg-f-04-0-UI 3', mRNA sequence.
ACCESSION
AW251306
VERSION
AW251306.1 GI:6594897
KEYWORDS
SOURCE
Rattus norvegicus (Norway rat)

```

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 511)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
PUBMED
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized ventricle at 13 dpc library cDNA library Preparation:
M.B. Soares Lab Clone distribution: clones will be available
through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source
1. .511
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJO-adg-f-04-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-BJO"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJO
library is a substracted library derived from the UI-R-AAL,
UI-R-AB1, UI-R-AD1, UI-R-AE1, UI-R-AF1, and
UI-R-AG1 libraries. These libraries represent tissues from
rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal
at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV
canal at 15 dpc, and ventricle at 13 dpc. The tag is a
string of 5-6 nucleotides present between the Not I site
and the oligo-dT track. The library was constructed as
described by Bonaldo, Lennon and Soares, Genome Research
6: 791-806, 1996.
TAG LIB=UI-R-BJO
TAG TISSUE=ventricle at 13 dpc
TAG_SEQ=CAGCGA"

BASE COUNT 127 a 113 g 151 t
ORIGIN

Query Match 7.3%; Score 171; DB 9; Length 511;
Best Local Similarity 78.4%; Pred. No. 3.3e-35;
Matches 261; Conservative 0; Mismatches 55; Indels 17; Gaps 4;

QY 879 ATGAAACGAGAGACCAACAGTATCCATTGATAGCGTCTCAGGACAGATGACAGAG 938
Db 331 ATGGAACGAGAGGCCACAGTTGTCCTTCGATAGTCTCAGGACGACGAG 272
QY 939 AGAACACTAGGAGGGGACCCACGAGGACCAAGGTATTAGTGTGTTGTTTCAGGCG 998
Db 271 ----CACTAGGAGGAGAACCCACGAGGA-----TATCAGTGTGCTGTTTCCAGGCG 221
QY 999 AATGCTCTTCTAGATCTAGAACCAATTTCTGTTGTAACAGCTGAAGTGGGT 1058
Db 220 AATGCTCTATTCGGAAGGTTCTAGAACACAGTTTTTCTGATGAAGAGCTGAAGTGGGT 161
QY 1059 GGGG---GTTCTTACCCCATGTTTCATGGAAGGCTGAGTGAGGAGACAGATATGATG 1115

Db 160 GGCGTAAGATAACACCGAGTTCACTGAAGGTGAGGAGGACGACATACGATG 101
QY 1116 GCCAGCATACAAACACACACCTTAATTAACACTTCCTCTCTTCTACTGACACCCC 1175
Db 100 GCCAGCAT-----ACAACATACACCTTAATTAATGCTTCCTCTCTGACTGACTCT 46
QY 1176 CTTCACTCTCTCTTTTCATAAAAAATAAAAAA 1208
Db 45 CCCTTCACTCTCTTTTCATAAAAAATAAAAAACA 13

RESULT 9
AI710529/c
LOCUS
DEFINITION
UI-R-AG1-aal-b-08-0-UI.s1 UI-R-AG1 Rattus norvegicus cDNA clone
UI-R-AG1-aal-b-08-0-UI 3', mRNA sequence.
AI710529
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 370)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
PUBMED
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
normalized ventricle at 13 dpc library cDNA library Preparation:
M.B. Soares Lab Clone distribution: clones will be available
through Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 1-43,
Seq primer: M13 Forward
POLYA=Yes

FEATURES
Location/Qualifiers
1. .370
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-AG1-aal-b-08-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-AG1"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-AG1
library is a normalized library constructed from 13 dpc
rat ventricle. The tag is a string of 6 nucleotides
present between the Not I site and the oligo-dT track.
The library was constructed as described by Bonaldo,
Lennon and Soares, Genome Research 6: 791-806, 1996.
Tissue provided by Jim Lin, Department of Biology,
University of Iowa.
TAG LIB=UI-R-AG1
TAG TISSUE=ventricle at 13 dpc
TAG_SEQ=CAGCGA"

```
BASE COUNT      75 a      82 c      80 g      132 t      1 others
ORIGIN

Query Match      7.1%; Score 167.8; DB 9; Length 370;
Best Local Similarity 77.8%; Pred. No. 2.2e-34;
Matches 259; Conservative 0; Mismatches 57; Indels 17; Gaps 4;

QY  879 ATGAAACGAGAGAGCAACAGTATCCATTGATAGCGTCTCAGGACAGATAGGACAGAG 938
    |||||
Db   348 ATGGAACGAGAGAGCAACAGTATCCATTGATAGCGTCTCAGGACAGATAGGACAGAG 289
    |||||
QY  939 AGAACACTAGGAGAGGGGAAACCCAGCAAGGATATTAGTGTTGGTTTTCAGGGC 998
    |||||
Db   288 ----CACTAGGAGAGAGAGAAACCCAGCAAGGA-----TATCAGTGTGCTGGTTCCAAGGCG 238
    |||||
QY  999 AATGCTCTTACTGAGATTCTAGAACACAAATTTCTGCTGTTGAACAGCTGAAGTGGGT 1058
    |||||
Db   237 AATGCTCTTCTCGAAGGTTCTAGAAACACAGTTTACTCGGTAAGAGCTGAAGTGGGT 178
    |||||
QY  1059 GGGG---GTCTTACCCCATGTTTCATGGAAGGTGAGTGAGGAGAGACAGATATATGATG 1115
    |||||
Db   177 GGGGTAAGATATACAAACCACTTCTACTGAAGGTGAGGAGGACGACAGACATACGATG 118
    |||||
QY  1116 GCCAGCATACAAACATACACACACCCCTAATTAACATTCCTCTTCTTACTGACACCCC 1175
    |||||
Db   117 GCCAGCAT-----ACAAACATAACACCCCTAATTAATGCTTCCCTCTGCTACTGACACTCT 63
    |||||
QY  1176 CTTCACTCTCTCTTTCATATAAAATAAAAAA 1208
    |||||
Db   62 CCCTTCACTCTCTTTCATATAAAATAAAACA 30
    |||||

RESULT 10
AUI139209
LOCUS      AUI139209 PLACE1 Homo sapiens cDNA clone PLACE1010155 5', mRNA
DEFINITION      sequence.
ACCESSION      AUI139209
VERSION        AUI139209.1 GI:11000730
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 735)
AUTHORS        Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
               Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
               Isogai,T.
TITLE          HRI human cDNA project
JOURNAL        Unpublished
COMMENT        Contact: Takao Isogai
               Genomics Laboratory
               Helix Research Institute
               1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
               Tel: 81-438-52-3975
               Fax: 81-438-52-3986
               Email: Genomics@hri.co.jp
               HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
               Research Institute; cDNA library construction: Department of
               Virology, Institute of Medical Science, University of Tokyo, and
               Helix Research Institute.
FEATURES       Location/Qualifiers
               1..735
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="PLACE1010155"
               /tissue_type="placenta"
               /clone_lib="PLACE1"
               /notes="Vector: pME18SFL3"
BASE COUNT      240 a      154 c      181 g      156 t      4 others
ORIGIN

Query Match      5.1%; Score 119.6; DB 10; Length 509;
Best Local Similarity 79.4%; Pred. No. 4e-21;
Matches 173; Conservative 0; Mismatches 32; Indels 2; Gaps 2;

QY  2105 TCATCTCCCTCTTGGGCTTCCAGACACTAGTCTGGAATGAAATTCACCTGCCTCTGA 2164
    |||||
Db   28  TCCCTCCCTCTTTCAGTCTCCAGACACTAGTCTGGAATGAAATTCACCTGCCTCTGA 87
    |||||
QY  2165 ATTGGCCACTGTGTGGGGCAGGGGTGTGACTTGGCTTCCAGGCTGGAGATTATCTCAC 2224
    |||||
Db   88  GTTGCTCTTAATGGGGGGGAGTGTACTTTCGGTTCAGAGTGGAGATTATCTCAC 147
    |||||
QY  2225 CCAGCCCTAGCTATATAA--CGGGCTGGTGTGAGGGGCTCCACAGGCCAGTTCAGGGG 2283
    |||||
Db   148 CCGGCCCCAGCTATATAAGCTGACCGGTGTGGAGGGGCCAGCAGGCCCACTCCAGGGA 207
    |||||
QY  2284 TTCATCCACAGAGAGAAAAACATAGA 2310
    |||||
Db   208 TTCTCTTC-CACGACAGAAAAACATACA 233
    |||||

RESULT 11
BE627514/c
LOCUS      BE627514
DEFINITION      509 bp mRNA linear EST 24-AUG-2000
               uu52e08.y1 Soares thymus 2NBMT Mus musculus cDNA clone
               IMAGE:3375590 5' similar to gb:U35933 Mouse erythrocyte protein 4.2
               (MOUSE);, mRNA sequence.
ACCESSION      BE627514
VERSION        BE627514.1 GI:9907936
KEYWORDS       EST.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 509)
AUTHORS        NCI-CCGAP Htcr://www.ncbi.nlm.nih.gov/ncicgap.
               National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL        Unpublished
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               This clone is available royalty-free through LNL; contact the
               IMAGE Consortium (info@image.lnl.gov) for further information.
               MGI:1085194
Seq primer: -40RP from Gibco
High quality sequence 465.
Location/Qualifiers
               1..509
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="C57BL/6J"
               /db_xref="taxon:10090"
               /clone="IMAGE:3375590"
               /sex="male"
               /tissue_type="Thymus"
               /dev_stage="4 weeks"
               /lab_host="DH10B"
               /clone_lib="Soares_thymus_2NBMT"
               /notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
               polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA
               was primed with a Not I - oligo(dT) primer [5',
               TGTTCACCAATCTGAAGTGGAGCGCGGTTTTTTTTTTTTTTTTTTTTTTT
               3']; double-stranded cDNA was ligated to Eco RI adaptors
               (Pharmacia), digested with Not I and cloned into the Not I
               and Eco RI sites of the modified pT7T3 vector. RNA
               provided by Dr. Bertrand Jordan. Library went through two
               rounds of normalization, and was constructed by Bento
               Soares and M. Fatima Bonaldo."
BASE COUNT      150 a      110 c      125 g      123 t      1 others
ORIGIN

Query Match      5.1%; Score 119.6; DB 10; Length 509;
Best Local Similarity 79.4%; Pred. No. 4e-21;
```

	Matches	154;	Conservative	0;	Mismatches	39;	Indels	1;	Gaps	1;
Qy	192	CCTGATCTTAGGGCAGCAGTTCTTCAACCTGGGGGCGCTCGACCCCTTTGGGGGAATCAAAAC	251							
Db	492	CCCTACCTTAGAGCAGCAGTTCTCAACCTGTGGGGCTCGACCCCTTTGGGGGTCGTATGA	433							
Qy	252	GACCCITTACAGGGGTCACATATCATCTATCCTATATGTCAGGTATTTACATACGATTC	311							
Db	432	TACT-TTCACAGGGGTAGGGTCATCAGGTATCCTCATATCAGATATTACATTAATGATTC	374							
Qy	312	GTAACAGTAGCAAAATTACAGGTATGAAATAGCAATGAAATAATTTTATGATTCAAGGTC	371							
Db	373	CTAACAGTAGCAAAATTTATAGTCTGAGTAGCAATGAATTAAGTTATGTTGGGGTC	314							
Qy	372	ACCACAACTAGG	385							
Db	313	ACCACAACTAGG	300							

RESULT 12					
AZ804735/c					
LOCUS	AZ804735	538 bp	DNA	linear	GSS 20-FEB-2001
DEFINITION	2M0065I113R Mouse 10kb plasmid UUC1M library Mus musculus genomic clone UUC2M0065I13 R, genomic survey sequence.				

```

FEATURES
    source
        1. 538
            Location/Qualifiers
                map_location="Mus musculus"
                mol_type="genomic DNA"
                strain="C57BL/6J"
                db_xref="taxon:10090"
                clone="UUGC2M0065113"
                sex="Male"
                lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                clone_lib="Mouse 10kb plasmid UUGC1M library"
                notes="vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose/gel

```

electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

164 a	127 c	112 g	135 t
-------	-------	-------	-------

BASE COUNT
ORIGIN

BASE COUNT	164 a	127 c	112 g	135 t
ORIGIN				

Query Match	4.9%	Score 116.2	DB 28	Length 538
Best Local Similarity	79.9%	Pred. No. 3.5e-20		
Matches 151: Conservative	0	Mismatches 33	Indels 5	Gaps 1

QY 197 TTCTAGGCAGCAGTTCTCAACCTGGGGCCCTCGACCCCTTTGGGGGAATCAACGACCC 256

DB	QY
390	257
TTCTAAGSCAATCTCTCTCAACCTACGGGTTTCGACCCCTTTGGG-----TATCAACCCCT	TTTACAGGGGTCACATATCATCTCTATATGTCTAGGTATTTCATTACGATTCGGTAAC

DB	335	TTACAGGGGTCACATATCACATATTCGCATGTCAGATATTTACATTACAGTTCAAC	276
OV	317	AGTAGCAAAATTACAGGTATGAAATAGCAATGAAATTAATTTTATGATTGAAGGTCAACC	376

RESULT 13

AQ480395/c

LOCUS AQ480395 522 bp DNA linear GSS 23-APR-1999

DEFINITION RPCI-11-236B22.TV RPCI-11 Homo sapiens genomic clone RPCI-11-236B22 , genomic survey sequence.

```

/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/notes="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCII Human Male BAC Library"
BASE COUNT 168 a 100 c 113 g 141 t
ORIGIN
Query Match 4.9%; Score 116; DB 28; Length 522;
Best Local Similarity 76.7%; Pred. No. 3.9e-20;
Matches 155; Conservative 0; Mismatches 45; Indels 2; Gaps 1;
QY 1540 TCTCTCTGTCATCACTCTGGCCCTTTGGGGT--AGATCCTCTGATTAGCTTCAGAT 1597
Db 202 TCTATTTCTTGACCACTCTGATCCATTTTGAGTAAATGCTCCAAATTATTATGCTGTT 143
QY 1598 TTAGAACACGGTAGCCTGTGGTGCACATAATPATGCCAGTGACACCATAGAGTCAAAAGT 1657
Db 142 TTAGAACACGGTAGCATGTCTGCTAAATTATGCGCAGTGACATCATATAAAGAAAGT 83
QY 1658 GCATTACTGAATGCTTTCAATTTCTTAATGCTGGTACGATGGCATGTCCACAGGCCAT 1717
Db 82 GCATTACTGAATGCTTTCAATTTCTTAATGCTGGTACGATGGCATGTCCATGGGGCCTA 23
QY 1718 TTTAGTCGACATCACTCCA 1739
Db 22 TTTAGCCCGACATCACTCCA 1

```

```

RESULT 14
LOCUS AZ114168 518 bp DNA linear GSS 12-MAY-2000
DEFINITION RPCI-23-449H23.TJ RPCI-23 Mus musculus genomic clone RPCI-23-449H23
, genomic survey sequence.
ACCESSION AZ114168
VERSION AZ114168.1 GI:7774139
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 518)
AUTHORS Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret
, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished
Other GSSs: RPCI-23-449H23.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(piet@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plate: 449 row: H column: 23
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. .518
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-449H23"
/sex="Female"
/lab_host="DH10B"

```

FEATURES
source

```

/clone_lib="RPCI-23"
/notes="Organ: Kidney/Brain; Vector: pBACE3.6; Site 1:
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACE3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT 158 a 94 c 108 g 158 t
ORIGIN

```

```

Query Match 4.9%; Score 114.6; DB 28; Length 518;
Best Local Similarity 67.2%; Pred. No. 9.4e-20;
Matches 162; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 145 ACAAGGTCTCATGGGAAATTCCTGGAGCTTCCTTTCCAGATCAGCTGATTCAGGG 204
Db 150 ACAGGCCATGATTTTGAAGGCTTAGTGTCTTTGATTTTAAACACTACTTTGACGACAG 209
QY 205 CAGCAGTTCTCAACCTGGGGCCTCGACCCCTTTGGGGAAATCAAAACGACCCCTTTACAGG 264
Db 210 CAGTGGTTCTCAACCTGTGGTGGGACCCCTGTGGGGGATTTACTGCTCTCCAGG 269
QY 265 GGTCAATATCATCTATCTATATGTCAGGTATTTACATTCAGATTTCGTAAACAGTAGCAA 324
Db 270 GGTCTCTTATCAATATTTCCACAGTCAGATATGATATTACAGTTTATAACAATGGCAA 329
QY 325 AATTACAGGTATGAAATAGCAATGAATAATTTTATGATTGAGTCCACCAACATGAG 384
Db 330 AATCCAGTTATGAGGGAGTAATGAATAATTTTATAGTTGAGGGTTCATCAACATGAG 389
QY 385 G 385
Db 390 G 390

```

```

RESULT 15
LOCUS AZ290439/596 bp DNA linear GSS 27-JUL-2000
DEFINITION RPCI-23-59G22.TVB RPCI-23 Mus musculus genomic clone RPCI-23-59G22,
genomic survey sequence.
ACCESSION AZ290439
VERSION AZ290439.1 GI:9532225
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 596)
AUTHORS Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret
, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished
Other GSSs: RPCI-23-59G22.TJB
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(piet@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plate: 59 row: G column: 22
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1. .596

```

FEATURES
source

```
/organism="Mus musculus"  
/mol_type="genomic DNA"  
/strain="C57BL/6J"  
/db_xref="taxon:10090"  
/clone="RPCI-23-59G22"  
/sex="Female"  
/lab_host="DH10B"  
/clone_lib="RPCI-23"  
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:  
EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or  
brain genomic DNA was isolated and partially digested  
with a combination of EcoRI and EcoRI Methyase. Size  
selected DNA was cloned into the pBACe3.6 vector at the  
EcoRI sites. The ligation products were transformed into  
DH10B electrocompetent cells (BRL Life Technologies). "  
BASE COUNT      174 a   136 c   119 g   167 t  
ORIGIN
```

```
Query Match      4.8%; Score 113; DB 28; Length 596;  
Best Local Similarity 78.3%; Pred. No. 2.7e-19;  
Matches 148; Conservative 0; Mismatches 40; Indels 1; Gaps 1;  
  
Qy 199 CTAGGGCAGCAGTCTCAACCTGGGGGCTCGACCCCTTTGGGGGAATCAACGACCCCTT 258  
    ||| ||||| || ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 467 CTATGGCAGTGGTCTCAACCTGGGTGGTGACCCCTTTGGGGG-TTCAACACCCCTTT 409  
  
Qy 259 TACAGGGTCACATATCATCTATCTATATGTCTCAGGTATTTACATTACGATTCGTAACAG 318  
    ||| ||||| ||||| ||| ||| ||||| ||||| ||||| ||||| |||||  
Db 408 CACAGGGGTCACATATCAGATTTTCTGCGTACCAGATATTTACATTATGATTCAGAACAG 349  
  
Qy 319 TAGCAAAATTACAGGTATGAAATAGCAATGAAATATTTTATGATTGAAGGTCAACACAA 378  
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 348 TAGCAAAATTCAGATTATGAAGTAGCAATGAAATATTTTGTGTGGGGGTCAGCACAA 289  
  
Qy 379 CATGAGGCC 387  
    |||||  
Db 288 CATGGGGAC 280
```

Search completed: November 14, 2003, 06:44:42
Job time : 4911.75 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 03:50:07 ; Search time 134.074 Seconds
(without alignments)
7762.738 Million cell updates/sec

Title: US-10-005-337A-1
Perfect score: 2358
Sequence: 1 ggaatcttcattttaaaca.....caggctggagccaccatgg 2358

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	88.6	3.8	48974	3	US-08-920-422-17
2	78	3.3	892	4	US-09-171-209-75
3	74.8	3.2	37950	3	US-09-338-907-183
4	74.8	3.2	37950	4	US-09-218-207-183
C 5	74.6	3.2	10614	1	US-08-135-511-35
C 6	74.6	3.2	10614	1	US-08-187-453-35
7	73.4	3.1	4164	4	US-08-882-164B-38
8	68.6	2.9	4072	3	US-09-272-496-7
C 9	68	2.9	16442	3	US-08-781-891-208
C 10	68	2.9	16442	4	US-09-618-166-208
C 11	65.6	2.8	29604	3	US-08-781-891-207
C 12	65.6	2.8	29604	4	US-09-618-166-207
C 13	65	2.8	6645	2	US-08-380-403A-4
C 14	65	2.8	6645	2	US-08-895-628-4
C 15	65	2.8	6645	4	US-08-895-810B-4
C 16	63	2.7	3450	2	US-08-378-617A-9
C 17	62	2.6	7208	3	US-09-166-186-107
C 18	62	2.6	7208	3	US-09-313-932-107
C 19	54	2.3	7218	1	US-08-232-463-14
20	50.6	2.1	5764	4	US-09-312-762A-8
21	47	2.0	26700	1	US-08-472-217-1
22	47	2.0	26700	2	US-08-488-199-5
23	47	2.0	26700	3	US-08-760-534A-1
24	47	2.0	26700	4	US-09-336-757-1
25	38.8	1.6	818	4	US-09-328-475C-157
26	38.6	1.6	2830	1	US-08-604-333-3
27	38.6	1.6	2830	3	US-09-110-618-3

28	38.6	1.6	2830	4	US-09-578-178-3	Sequence 3, Appli
29	38.6	1.6	2830	4	US-09-577-806-3	Sequence 3, Appli
30	37.2	1.6	7218	1	US-08-232-463-14	Sequence 14, Appli
31	36.8	1.6	277	3	US-09-007-005-3	Sequence 3, Appli
32	36.8	1.6	277	3	US-09-244-796-3	Sequence 3, Appli
33	36.6	1.6	1664976	4	US-08-916-421B-1	Sequence 1, Appli
34	36.2	1.5	1209	4	US-08-462-509B-1	Sequence 1, Appli
35	36.2	1.5	1209	5	PCT-US95-05616-1	Sequence 1, Appli
36	36.2	1.5	1564	2	US-08-846-705-4	Sequence 4, Appli
37	36.2	1.5	1564	3	US-08-846-704-1	Sequence 1, Appli
38	36.2	1.5	1564	3	US-08-846-704-3	Sequence 3, Appli
39	35	1.5	9573	4	US-09-220-132-168	Sequence 168, App
C 40	35	1.5	80246	3	US-09-078-294-4	Sequence 4, Appli
C 41	35	1.5	80595	3	US-09-078-294-3	Sequence 3, Appli
42	34.8	1.5	2088	2	US-08-602-264A-1	Sequence 1, Appli
43	34.8	1.5	2088	3	US-08-461-018A-1	Sequence 1, Appli
44	34.8	1.5	2088	3	US-09-216-958-1	Sequence 1, Appli
C 45	34.4	1.5	1843	4	US-09-622-540A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-920-422-17/c
; Sequence 17, Application US/08920422A
; Patent No. 6255473
; GENERAL INFORMATION:
; APPLICANT: Vitek, Michael P.
; APPLICANT: Mitsuda, No. 6255473iaki
; TITLE OF INVENTION: Presenilin-1 Gene Promoter
; FILE REFERENCE: VITEKPRESENTIN
; CURRENT APPLICATION NUMBER: US/08/920,422A
; CURRENT FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 17
; LENGTH: 48974
; TYPE: DNA
; ORGANISM: Mus musculus
US-08-920-422-17

Query Match 3.8%; Score 88.6; DB 3; Length 48974;
Best Local Similarity 69.2%; Pred. No. 3.9e-16;
Matches 135; Conservative 0; Mismatches 59; Indels 1; Gaps 1;
QY 191 GCTGATTCCTAGGCGACAGTCTCTCAACCTGGGGGCTCGACCCCTTTGGGGAATCAAA 250
DB 15385 GCTACACTCTAGATCGTGTCTCAAGCTGGGGGTCGCCGCCCTTTGGGGTTGGACA 15326
QY 251 CGACCCCTTTACAGGGGTACATATCATCTATCTATATGTCAGGTATTTACATACGATT 310
DB 15325 A-CCTTTTCACAGGGGTACACATCAGATCTCGCATATTTATGTTATGATT 15267
QY 311 CGTAAACAGTAGCAAAATTTACAGGTATGAAATACAAATGAATATTTATGATTGAAGGT 370
DB 15266 CATAACAGCAGACAGATTACAGTTAGGAAGTAAATGAATATGTTATGTTGGGC 15207
QY 371 CACCAACAATGAGG 385
DB 15206 CACCATGACATGAAG 15192

RESULT 2
US-09-171-209-75
; Sequence 75, Application US/09171209
; Patent No. 6448000
; GENERAL INFORMATION:
; APPLICANT: VANDERBILT UNIVERSITY
; 305 Kirkland Hall
; Nashville, TN 37240
; TITLE OF INVENTION: MAMMALIAN GENES INVOLVED IN VIRAL

```
/ INFECTION
/ NUMBER OF SEQUENCES: 83
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Needle & Rosenberg, P.C.
/ STREET: 127 Peachtree Street, Suite 1200
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: USA
/ ZIP: 30303-1811
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA: US/09/171,209
/ FILING DATE: 08-Mar-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US97/06067
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Selby, Elizabeth
/ REGISTRATION NUMBER: 38,298
/ REFERENCE/DOCKET NUMBER: 22000.0061/P
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404 688 0770
/ TELEFAX: 404 688 9880
/ INFORMATION FOR SEQ ID NO: 75:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 892 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ SEQUENCE DESCRIPTION: SEQ ID NO: 75:
US-09-171-209-75

Query Match 3.3%; Score 78; DB 4; Length 892;
Best Local Similarity 70.5%; Pred. No. 6.9e-14;
Matches 148; Conservative 0; Mismatches 55; Indels 7; Gaps 3;

QY 177 CTCCTTCAGGATCAGCCTCATCTTAGGGCAGCAGTCTCACTGGGCGCCCTCGACCCCT 236
Db 119 CAGGTACATAGATAGTCAAAATCTAGACACTGTCTTCTATCTCTGTGAGTTGCAACCCCT 178
QY 237 TTGGG---GGAATCAACGACCCCT-TTACAGGGGTCACATATCATCTATCTATATGTCA 292
Db 179 TTGGGAGTGGGTCAAATGACCCCTATCACAGGGGTCTCAAATGAGATATCTCTGCATATCA 238
QY 293 GGTATTACATTACGATTCTGTACAGTAGCAAAATTTACAGGTATGAAATAGCAATGAAT 352
Db 239 AATATTACATTATGATTCATAGTAGTACAGAAATTTACAGTTATGAGTTTACA---AAAT 295
QY 353 AATTTTATGATTGAAGTGCACCAACATG 382
Db 296 AATTTTATAGCTGAGAGTCACCAACATG 325

RESULT 3
US-09-338-907-183
/ Sequence 183, Application US/09338907
/ Patent No. 6265546
/ GENERAL INFORMATION:
/ APPLICANT: Blumenfeld, Marta
/ APPLICANT: Cohen, Daniel
/ APPLICANT: Ilya, Chumakov
/ APPLICANT: Bougueleret, Lydie
/ TITLE OF INVENTION: PROSTATE CANCER GENE
/ FILE REFERENCE: GENSET.18C1PC
/ CURRENT APPLICATION NUMBER: US/09/338,907
/ CURRENT FILING DATE: 1999-06-23
/ EARLIER APPLICATION NUMBER: 08/996,306

/ NUMBER OF SEQUENCES: 83
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Needle & Rosenberg, P.C.
/ STREET: 127 Peachtree Street, Suite 1200
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: USA
/ ZIP: 30303-1811
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA: US/09/171,209
/ FILING DATE: 08-Mar-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US97/06067
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Selby, Elizabeth
/ REGISTRATION NUMBER: 38,298
/ REFERENCE/DOCKET NUMBER: 22000.0061/P
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404 688 0770
/ TELEFAX: 404 688 9880
/ INFORMATION FOR SEQ ID NO: 75:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 892 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ SEQUENCE DESCRIPTION: SEQ ID NO: 75:
US-09-171-209-75

Query Match 3.3%; Score 74.8; DB 3; Length 37950;
Best Local Similarity 80.0%; Pred. No. 6.5e-12;
Matches 88; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 276 ATCTATCCTATATGTCAGGTATTTACATTACAGTTCGTTACAGTACCAAAATTCAGGTA 335
Db 9570 ACCTATCCTGCTTATCAGATAGTTACATTATGAATTTGTAACAGCAGCAAAATTCACAGTTA 9629
QY 336 TCAATAGCAATGAATATTTTATGATTGAAGGTCCACCAACATGAGG 385
Db 9630 CGCAATATCAACAAATAATTTTATGTTGAGGGTCACCATACGTGAGG 9679

RESULT 4
US-09-218-207-183
/ Sequence 183, Application US/09218207
/ Patent No. 6346381
/ GENERAL INFORMATION:
/ APPLICANT: Blumenfeld, Marta
/ APPLICANT: Cohen, Daniel
/ APPLICANT: Ilya, Chumakov
/ APPLICANT: Bougueleret, Lydie
/ TITLE OF INVENTION: Prostate cancer gene
/ FILE REFERENCE: GENSET.018CPI
/ CURRENT APPLICATION NUMBER: US/09/218,207
/ CURRENT FILING DATE: 1998-12-22
/ EARLIER APPLICATION NUMBER: 08/996,306
/ EARLIER FILING DATE: 1997-12-22
/ EARLIER APPLICATION NUMBER: 60/099,658
/ EARLIER FILING DATE: 1998-09-09
/ NUMBER OF SEQ ID NOS: 578
/ SOFTWARE: Patent.pm

/ EARLIER FILING DATE: 1997-12-22
/ EARLIER APPLICATION NUMBER: 60/099,658
/ EARLIER FILING DATE: 1998-09-09
/ NUMBER OF SEQ ID NOS: 578
/ SOFTWARE: Patent.pm
/ SEQ ID NO 183
/ LENGTH: 37950
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 5259..5328
/ OTHER INFORMATION: exon2
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 12675..12791
/ OTHER INFORMATION: exon3
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 14621..14710
/ OTHER INFORMATION: exon4
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 19822..19912
/ OTHER INFORMATION: exon5
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 21789..21950
/ OTHER INFORMATION: exon6
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 23387..23510
/ OTHER INFORMATION: exon7
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 25520..26016
/ OTHER INFORMATION: exon8
/ OTHER INFORMATION:
US-09-338-907-183
```

```

; SEQ_ID NO 183
; LENGTH: 37950
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5259..5328
; OTHER INFORMATION: exon2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 12675..12791
; OTHER INFORMATION: exon3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 14621..14710
; OTHER INFORMATION: exon4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 19822..19912
; OTHER INFORMATION: exon5
; FEATURE:
; NAME/KEY: exon
; LOCATION: 21789..21950
; OTHER INFORMATION: exon6
; FEATURE:
; NAME/KEY: exon
; LOCATION: 23387..23510
; OTHER INFORMATION: exon7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25520..26016
; OTHER INFORMATION: exon8
US-09-218-207-183

```

		Query Match	3.2%	Score 74.8;	DB 4;	Length 37950;	
		Best Local Similarity	80.0%;	Pred. No. 6.5e-12;			
		Matches	89;	Conservative	0;	Mismatches	22;
						Indels	0;
						Gaps	0;
QY	276	ATCTTCTCTATATGTCAGGTATTACATTACGATTTCGTAACAGTAGCAAAATAATCACAGTTA	335				
Dd	9570	ACCTATCCTGCTTATCATAGATAGTTACATTATGAATTGTAACAGCAGCAAAATCACAGTTA	9629				
QY	336	TGAAATAGCAATGAAATAATTTTTATGATTGAAGGTACCACAACATCAGG	385				
Dd	9630	CGCAAATATCAACAAAATAATTTTTATGGTTCAGGGTCCACCTAACGTCAGG	9679				

```

RESULT 5
US-08-135-511-35/c
; Sequence 35, Application US/08135511
; Patent No. 5558999
; GENERAL INFORMATION:
; APPLICANT: Chiang, John
; TITLE OF INVENTION: Cholesterol 7a-Hydroxylase Gene
; TITLE OF INVENTION: Regulatory Elements and Methods for Using Them
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/135,511
; FILING DATE: 13-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SANDERCOCK, Colin G.

```

```

, , REGISTRATION NUMBER: 31,298
, , REFERENCE/DOCKET NUMBER: 18748/175
, , TELECOMMUNICATION INFORMATION:
, , TELEPHONE: (202)672-5300
, , TELEFAX: (202)672-5399
, , TELEX: 904136
, , INFORMATION FOR SEQ ID NO: 35:
, , SEQUENCE CHARACTERISTICS:
, , LENGTH: 10614 base pairs
, , TYPE: nucleic acid
, , STRANDEDNESS: single
, , TOPOLOGY: linear
, , MOLECULE TYPE: DNA (genomic)
US-08-135-511-35

```

Query Match	3.2%;	Score 74.6;	DB 1;	Length 10614;
Best Local Similarity	64.0%;	Pred. No. 3.5e-12;		
Matches 130;	Conservative 0;	Mismatches 69;	Indels 4;	Gaps 1;
QY	195	GATTTCTAGGGCAGCAGCTTCTCAACCTGTGGGGCTCGACCCCTTTGGGG---	CAATCAAA	250
DB	7215	GRATCTAAGTCAGTGTCTTCTCAACCTGTGGATCACAAACCCATGTAGGGGTAGAGTCAAA		7156
QY	251	CGACCCCTTTACAGGGGTCAATCATCTATCTCTATATGTACGGTATTTACATTACGATT		310
DB	7155	TGACCCCTTTTCACAGGGGTCACTTAACACCATCAGAAAACACAGATATTTGCTTTATGATT		7096
QY	311	CGTAAACAGTAGCAAAATTTACAGTATGAAATAGCAATGAAATAAATTTATGATTGAAGGT		370
DB	7095	CATGACAGAAAGCAAAATTTATGTTGTGAAGTAGCAATGAANAATATTTACAGTTGGGGG		7036
QY	371	CACCAACAACATGAGCGCCACACA	393	
DB	7035	TCACACAAACATCAGGAAGGTAGA	7013	

RESULT 6
US-08-187-453-35/c
; Sequence 35, Application US/08187453
; Patent No. 5753431
; GENERAL INFORMATION:
; APPLICANT: Chiang, John
; TITLE OF INVENTION: Cholesterol 7a-Hydroxylase Gene
; TITLE OF INVENTION: Regulatory Elements and Transcription Factors
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/187,453
; FILING DATE: 28-JAN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/135,488
; FILING DATE: 13-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/135,511
; FILING DATE: 13-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/135,510
; FILING DATE: 13-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SANDERCOCK, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 18748/188

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 10614 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-187-453-35

Query Match 3.2%; Score 74.6; DB 1; Length 10614;
Best Local Similarity 64.0%; Pred. No. 3.5e-12;
Matches 130; Conservative 0; Mismatches 69; Indels 4; Gaps 1;

QY 195 GATTCAGGCGCAGCAGTCTCTCAACCTGGGGGCTCGACCCCTTTGGGG---GAATCAAA 250
Db 7215 GAATCTAAGTCAGTGTCTCAACCTGTGGATCACAAACCATGTAGGGGTAGAGTCAAA 7156

QY 251 CGACCCCTTTACAGGGGTCAATATCATCTATCTATATGTCTAGGTATTTACATTTAGATT 310
Db 7155 TGACCCCTTTACAGGGGTCACTTAAGACCATCAGAAAACACAGATATTTGCTTTATGATT 7096

QY 311 CGTAACAGTAGCAAAATTTACAGGTATGAATAGCAATGAATAATTTATGATTGAAGCT 370
Db 7095 CATGACAGAGCAAAATTTATAGTTGTGAAGTAGCAATGAATAATTTACAGTTGGGGG 7036

QY 371 CACCACAACATAGGCGGCCACA 393
Db 7035 TCACACAACATCAGGAAGGTAGA 7013

RESULT 7
US-08-882-164D-38
Sequence 38, Application US/08882164D
Patent No. 6306624
GENERAL INFORMATION:
APPLICANT: Petkovich, P. Martin, White, Jay A.,
APPLICANT: Beckett, Barbara R., Jones, Glenville
TITLE OF INVENTION: Retinoid Metabolizing Protein
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Blake, Cassels & Graydon
STREET: Box 25, Commerce Court West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5L 1A9
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
COMPUTER: COMPAQ, IBM PC compatible
OPERATING SYSTEM: MS-DOS 5.1
SOFTWARE: WORD PERFECT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882,164D
FILING DATE: June 25, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,546
FILING DATE: June 21, 1996
APPLICATION NUMBER: 08/724,466
FILING DATE: October 1, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunt, John C.
REGISTRATION NUMBER: 36,424
REFERENCE/DOCKET NUMBER: 50767/00010
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 863-4344
TELEFAX: (416) 863-2653
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 4164 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-882-164D-38

Query Match 3.1%; Score 73.4; DB 4; Length 4164;
Best Local Similarity 70.2%; Pred. No. 4.7e-12;
Matches 127; Conservative 0; Mismatches 51; Indels 3; Gaps 2;

QY 206 AGCAGTTCTCAACCTGGGGCTCGACCCCTTTGGGGGAATCAAGGACCTTTACAGGG 265
Db 1252 AGAGCTTTTCAACCTGTGGGTCTGACCCCTTTCACGAGCCCAACAAACCCCTTTCAGAAGG 1311

QY 266 GTCACATATCATCTATCTATATGTCAAGTATTTTACATTCGTAACAGTACGACAAA 325
Db 1312 GTCGCTTAAGAGC--ATCTGCAATATCCGATATTTATCATCAAGAAACATACAGTACAAA 1369

QY 326 ATTACAGGTATGAAATAGCAATGAA--ATAATTTTATGATTGAAGGTCCACCAACATGAG 384
Db 1370 ATTACCGTTATGAGTAGCAACAAAGATAATTTTATCGTTGGGGTCCACCAACACGAG 1429

QY 385 G 385
Db 1430 G 1430

RESULT 8
US-09-272-496-7
Sequence 7, Application US/09272496
Patent No. 6245966
GENERAL INFORMATION:
APPLICANT: DeGregori, James
TITLE OF INVENTION: Adenoviral mediated gene transfer into lymphocytes
FILE REFERENCE: 90-98
CURRENT APPLICATION NUMBER: US/09/272,496
CURRENT FILING DATE: 1999-03-19
EARLIER APPLICATION NUMBER: US 60/092782
EARLIER FILING DATE: 1998-07-14
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 7
LENGTH: 4072
TYPE: DNA
ORGANISM: Mus musculus
US-09-272-496-7

Query Match 2.9%; Score 68.6; DB 3; Length 4072;
Best Local Similarity 69.5%; Pred. No. 1.4e-10;
Matches 141; Conservative 0; Mismatches 49; Indels 13; Gaps 3;

QY 195 GATTCAGGCGCAGCAGTCTCTCAACCTGGGGGCTCGACCCCTTTGGGGGAATCAACGAC 254
Db 50 GACCGGTAGACGACAGGTTCTCAATCTGTGGTCCGCAACCCCTTT-GGGCATCCAAAAGCA 108

QY 255 CTTTACAGGGGTCAATATCATCTATCTATATGTCAAGTATTTTACA-----T 303
Db 109 CTTTACAGGGTTCATATCATATCCAGTATATCAGATATTTATTTAAATATT 168

QY 304 TACGATTCGTAAACAGTAGCAAAATTTACAGGTATGAATAGCAATG-AAATAATTTTATGA 362
Db 169 TAAAAGTCGTAGCAGTGGCAAGATTACGGTTTACAAAGTAGCAACGAAATAATTTTATGCG 228

QY 363 TTGAAGGTCAACCAACATGAGG 385
Db 229 TTGGGAGTCATCAGCATGAGG 251

RESULT 9
US-08-781-891-208/c
Sequence 208, Application US/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui

APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 208:
SEQUENCE CHARACTERISTICS:
LENGTH: 16442 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-208
Query Match 2.9%; Score 68; DB 3; Length 16442;
Best Local Similarity 66.2%; Pred. No. 5e-10;
Matches 98; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 238 TGGGGGAATCAACAGCACCCCTTTACAGGGGTCAACATATCTATCTATATGTCAGGTAT 297
Db 6043 TGATTGATCAATGGCTCTTTACAGGGTTTCAAAATCAACATCTCTGTATATCAGATAG 5984
QY 298 TTACATTACGATTCGTAACAGTAGCAAAATTCAGGTATGAAATAGCAATGAATATTT 357
Db 5983 TTACATCACAATTCATAACGGTGGCAAAATTCAGTCAAGAGTAGCAAGGAAATAATG 5924
QY 358 TATGATTGAAGGTCAACACACATGAGG 385
Db 5923 TCATGGTTGGTGCCATCAGACGAGG 5896
RESULT 10
US-09-618-166-208/c
Sequence 208, Application US/09618166
Patent No. 6583112
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
Yu, Chang-En
Oshima, Junko
Mulligan, John T.
Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington

COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/618,166
FILING DATE: 17-Jul-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052.419C1
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 208:
SEQUENCE CHARACTERISTICS:
LENGTH: 16442 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-618-166-208
Query Match 2.9%; Score 68; DB 4; Length 16442;
Best Local Similarity 66.2%; Pred. No. 5e-10;
Matches 98; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 238 TGGGGGAATCAACAGCACCCCTTTACAGGGGTCAACATATCTATCTATATGTCAGGTAT 297
Db 6043 TGATTGATCAATGGCTCTTTACAGGGTTTCAAAATCAACATCTCTGTATATCAGATAG 5984
QY 298 TTACATTACGATTCGTAACAGTAGCAAAATTCAGGTATGAAATAGCAATGAATATTT 357
Db 5983 TTACATCACAATTCATAACGGTGGCAAAATTCAGTCAAGAGTAGCAAGGAAATAATG 5924
QY 358 TATGATTGAAGGTCAACACACATGAGG 385
Db 5923 TCATGGTTGGTGCCATCAGACGAGG 5896
RESULT 11
US-08-781-891-207
Sequence 207, Application US/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800

```
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 207:
SEQUENCE CHARACTERISTICS:
LENGTH: 29604 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-207

Query Match      2.8%; Score 65.6; DB 3; Length 29604;
Best Local Similarity 76.9%; Pred. No. 4e-09;
Matches 80; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 262 AGGGGTACATATCTCTATCTATGTCAGGTATTTACATTACAGTTCGTAACAGTAG 321
Db 18748 AGCGACCCCTTCACAGATATCTGAATATCAGGTATTTACATCGTGATTCATAGCAGTAA 18807

QY 322 CAAAATTACAGTATGAATAGCAATGAATTAATTTATGATTG 365
Db 18808 CAAAATTACAGTATGAATAGCAATGAATTAATTTATGATTG 18851

RESULT 12
US-09-618-166-207
; Sequence 207, Application US/09618166
; Patent No. 6583112
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; Yu, Chang-En
; Oshima, Junko
; Mulligan, John T.
; Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESS: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/618,166
FILING DATE: 17-Jul-2000
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052.419C1
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 207:
SEQUENCE CHARACTERISTICS:
LENGTH: 29604 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 207:
US-09-618-166-207
```

```
Query Match      2.8%; Score 65.6; DB 4; Length 29604;
Best Local Similarity 76.9%; Pred. No. 4e-09;
Matches 80; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 262 AGGGGTACATATCTCTATCTATGTCAGGTATTTACATTACAGTTCGTAACAGTAG 321
Db 18748 AGCGACCCCTTCACAGATATCTGAATATCAGGTATTTACATCGTGATTCATAGCAGTAA 18807

QY 322 CAAAATTACAGTATGAATAGCAATGAATTAATTTATGATTG 365
Db 18808 CAAAATTACAGTATGAATAGCAATGAATTAATTTATGATTG 18851

RESULT 13
US-08-380-403A-4/C
; Sequence 4, Application US/08380403A
; Patent No. 5831024
; GENERAL INFORMATION:
; APPLICANT: MINATO, Nagahito
; APPLICANT: HATTORI, Masakazu
; APPLICANT: HIROSHI, Kubota
; APPLICANT: MASATSUGU, Naeda
; TITLE OF INVENTION: SPA-1 PROTEIN AND GENE CODING THEREFOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/380,403A
; FILING DATE: 30-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/325,909
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-279712
; FILING DATE: 20-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-139513
; FILING DATE: 30-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/128/AAOK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6645 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Join(904..1015, 1356..1459, 1726..1883, 2009
; LOCATION: ..2618, 2890..3164, 4291..4509, 4598..4709, 4795
; LOCATION: ..4903, 5017..5117, 5200..5255, 5447..5525, 5598
; LOCATION: ..5741)
; US-08-380-403A-4

Query Match      2.8%; Score 65; DB 2; Length 6645;
```



```
; NAME/KEY: CDS
; LOCATION: join(904..1015, 1356..1459, 1726..1883, 2009
; :2618, 2890..3164, 4291..4509, 4598..4709, 4795..4903,
; 5017..5117, 5200..5255, 5447..5525, 5598..5741)
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-895-810D-4

Query Match      2.8%; Score 65; DB 4; Length 6645;
Best Local Similarity 72.3%; Pred. No. 2.5e-09;
Matches 112; Conservative 0; Mismatches 40; Indels 3; Gaps 2;

QY 233 CCCTTTGGGGGAATCAAAAGACCCCTTTACAGGGGTCACAT--ATCATCTATCTATATGT 290
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 554 CCCTTCTGAGGGTCTAAACGGCCCAATTCACAGGGTCCCATGCAGCAGACCCCTGTGTGT 495

QY 291 CAGGTATTTACATTACGATTCGTACAGTAGCAAAATTCAGGTATGAATAATAGCAATGAA 350
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 494 CAGGTGTTTTATGGTACAATTCCTAAGGGTAGCAGAATTACAGTTATGAAGTGCAATGAA 435

QY 351 ATAATTTTATGATGAAGGTACCAACAACATGAGG 385
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 434 AATAATTTTGA--TCAGGGTCACCACCATGTGG 401
```

Search completed: November 14, 2003, 06:49:00
Job time : 138.074 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 03:51:42 ; Search time 698.036 Seconds
(without alignments)
11040.154 Million cell updates/sec

Title: US-10-005-337A-1
Perfect score: 2358
Sequence: 1 99atcttctcatgtttaaca.....caggctggaggccaccatgg 2358

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2169961 seqs, 1634102185 residues
Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2358	100.0	2358	14	US-10-005-337A-1
2	367.8	15.6	2074	14	Sequence 1, Appli
3	100.4	4.3	90650	14	Sequence 2, Appli
4	88.8	3.8	659158	10	Sequence 80, Appl
5	84.2	3.6	106664	14	Sequence 20, Appl
6	82	3.5	659158	10	Sequence 97, Appl
7	80.6	3.4	889	12	Sequence 20, Appl
8	80.4	3.4	10917	12	Sequence 156, Appl
9	78.2	3.3	185548	14	Sequence 1, Appli
10	78	3.3	892	14	Sequence 62, Appl
11	76.8	3.3	249487	13	Sequence 175, Appl
12	75.8	3.2	335	10	Sequence 3, Appli
13	75.6	3.2	5990	10	Sequence 91, Appl
14	74.8	3.2	37950	10	Sequence 917-800A-477
15	74.8	3.2	37950	10	Sequence 183, Appl
16	71.6	3.0	171936	12	Sequence 183, Appl
					Sequence 24, Appli

C	17	71.6	3.0	171936	12	US-10-265-071-24	Sequence 24, Appl
	18	69.4	2.9	1889	10	US-09-974-298-182	Sequence 182, App
	19	69.4	2.9	1889	12	US-10-240-965-158	Sequence 158, App
	20	69.4	2.9	1889	12	US-10-252-157-421	Sequence 421, App
C	21	68.8	2.9	106664	14	US-10-175-523-97	Sequence 97, Appl
C	22	68	2.9	185548	14	US-10-175-523-62	Sequence 62, Appl
	23	67.8	2.9	1325	14	US-10-084-817-217	Sequence 217, App
C	24	64	2.7	85548	14	US-10-175-523-75	Sequence 75, Appl
	25	63.8	2.7	37051	12	US-10-004-113-55	Sequence 55, Appl
	26	63.2	2.7	184	9	US-09-923-876-5218	Sequence 5218, Ap
	27	63	2.7	3240	12	US-10-125-994A-25	Sequence 25, Appl
	28	63	2.7	3537	12	US-10-125-994A-26	Sequence 26, Appl
C	29	62	2.6	7208	11	US-09-824-322B-107	Sequence 107, App
C	30	60.6	2.6	67191	12	US-10-105-613-1	Sequence 1, Appli
	31	60.4	2.6	696	9	US-09-728-446-115	Sequence 115, Appl
C	32	60	2.5	32069	12	US-10-004-113-7	Sequence 7, Appli
C	33	59.8	2.5	347	9	US-09-728-446-339	Sequence 339, Appl
	34	59.6	2.5	4109	11	US-09-866-050A-567	Sequence 567, App
	35	59.6	2.5	4109	14	US-10-152-661-567	Sequence 567, App
	36	59.4	2.5	520	10	US-09-867-701-1380	Sequence 1380, Ap
C	37	57.6	2.4	6000	9	US-09-738-968-44	Sequence 44, Appl
C	38	57.4	2.4	893	13	US-10-044-592-3	Sequence 3, Appli
C	39	57.4	2.4	955	13	US-10-044-592-93	Sequence 93, Appl
C	40	57.4	2.4	1574	13	US-10-044-592-95	Sequence 95, Appl
C	41	56.6	2.4	1235	10	US-09-809-545A-15	Sequence 15, Appl
C	42	56.6	2.4	2120	9	US-09-801-574-29	Sequence 29, Appl
	43	56.2	2.4	74868	14	US-10-175-523-67	Sequence 67, Appl
	44	55.6	2.4	1470	14	US-10-208-304-6	Sequence 6, Appli
C	45	55.6	2.4	173808	13	US-10-003-806-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-10-005-337A-1
; Sequence 1, Application US/10005337A
; Publication No. US20030039984A1
; GENERAL INFORMATION:
; APPLICANT: BENOIT, Patrick
; APPLICANT: SCHWARTZ, Bertrand
; APPLICANT: BRANELLEC, Didier
; APPLICANT: CHEN, Kenneth R.
; TITLE OF INVENTION: SEQUENCES UPSTREAM OF THE CARP GENE, VECTORS CONTAINING
; TITLE OF INVENTION: THEM AND USES THEREOF
; FILE REFERENCE: 03806.0530-00000
; CURRENT APPLICATION NUMBER: US/10/005,337A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/251,582
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2358
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-005-337A-1

Query Match	100.0%;	Score	2358;	DB	14;	Length	2358;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	2358;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	GGATCTTTTCATGTTTAAACAATCAACCCCAAGGGGAACAGCCCTGCCTGACAGT	60				
Db	1	GGATCTTTTCATGTTTAAACAATCAACCCCAAGGGGAACAGCCCTGCCTGACAGT	60				
Qy	61	GGCTTTGCCACCATGAATACTTCTAGTCTAGTCCGTTTGTGAACTCAGCCCATCCCA	120				
Db	61	GGCTTTGCCACCATGAATACTTCTAGTCTAGTCCGTTTGTGAACTCAGCCCATCCCA	120				
Qy	121	ACACTTCTGCAAGCCCATCCTCTACAAGTGCTCATTTGGGAATTTCTCGAGCTTCT	180				
Db	121	ACACTTCTGCAAGCCCATCCTCTACAAGTGCTCATTTGGGAATTTCTCGAGCTTCT	180				

181 TTCAGGATCAGCCTGATTCTTAGGCGAGCAGTTCTCAACCTGGGGCCTCGACCCCTTTGG 240
181 TTCAGGATCAGCCTGATTCTTAGGCGAGCAGTTCTCAACCTGGGGCCTCGACCCCTTTGG 240
241 GGGAAATCAACAGCCCTTTACAGGGGTGCATATCATCTATCTATATGTTCAGGTATTTA 300
241 GGGAAATCAACAGCCCTTTACAGGGGTGCATATCATCTATCTATATGTTCAGGTATTTA 300
301 CATTTAGATTCGTAAACAGTAGCAAAATTAAGGTATGAATAGCAATGAATAATTTTAT 360
301 CATTTAGATTCGTAAACAGTAGCAAAATTAAGGTATGAATAGCAATGAATAATTTTAT 360
361 GATTGAAGGTCAACCAATGAGCGCCGACACTGTTCTAGAGAAATATCACTGGGTG 420
361 GATTGAAGGTCAACCAATGAGCGCCGACACTGTTCTAGAGAAATATCACTGGGTG 420
421 GGGAAAGGTTTGGAAAGCCTTTCTGTCCATTCTTCAATCTTCAAAAGTGAATGTTTCA 480
421 GGGAAAGGTTTGGAAAGCCTTTCTGTCCATTCTTCAATCTTCAAAAGTGAATGTTTCA 480
481 GAAAGCCTTTCAAGTCTTCTGTGGGCTCTTAGTAAGTCTGAGTAGGAACCTGATGTAT 540
481 GAAAGCCTTTCAAGTCTTCTGTGGGCTCTTAGTAAGTCTGAGTAGGAACCTGATGTAT 540
541 CAGGTCTGCTTCTTATAGGCGGAGCAAGCGCATCGTGGGTGAGCGAAGCGCAACCT 600
541 CAGGTCTGCTTCTTATAGGCGGAGCAAGCGCATCGTGGGTGAGCGAAGCGCAACCT 600
601 CACCTTCTAGCTCTGATCCATAGCAAGTGAAGTCTTCTGTCTAGGTGTCATCT 660
601 CACCTTCTAGCTCTGATCCATAGCAAGTGAAGTCTTCTGTCTAGGTGTCATCT 660
661 CTGTGAATCGAGATCCTTGGCCTTGCTTGAATTAGGGAGCGCAAAATATCTCAGAGATTC 720
661 CTGTGAATCGAGATCCTTGGCCTTGCTTGAATTAGGGAGCGCAAAATATCTCAGAGATTC 720
721 AAGACTGCTCAGAGCGCCAGAGTCTTCTCAAGGAAAGGTCATCACTCTCAGCCCTCC 780
721 AAGACTGCTCAGAGCGCCAGAGTCTTCTCAAGGAAAGGTCATCACTCTCAGCCCTCC 780
781 TTAGCTCTCAGTCTGAGTCTGAAACAAACCGGCCACAGGAATGAGAAAGTCTGCAATAGCTG 840
781 TTAGCTCTGAGTCTGAGTCTGAAACAAACCGGCCACAGGAATGAGAAAGTCTGCAATAGCTG 840
841 CTTGTCACTTCAAGAGGTCAAGAAATAGTGTAAACCATGAAACGAGAAAGCAACACAG 900
841 CTTGTCACTTCAAGAGGTCAAGAAATAGTGTAAACCATGAAACGAGAAAGCAACACAG 900
901 TTATCCATTGATAGCGTCTCAGGACAGATAGGACAGAGACACTAGGAGGGGAACC 960
901 TTATCCATTGATAGCGTCTCAGGACAGATAGGACAGAGACACTAGGAGGGGAACC 960
961 CACGAAGGACAGGATATTAGTGTGTTTTCAGGGCAATGTCTTGTACTGAAGATTCT 1020
961 CACGAAGGACAGGATATTAGTGTGTTTTCAGGGCAATGTCTTGTACTGAAGATTCT 1020
1021 AGAAAACAAATTTGCTGTTGAAACAGTGAAGTGGGGTGGGGTCTTACCCCATTTCA 1080
1021 AGAAAACAAATTTGCTGTTGAAACAGTGAAGTGGGGTGGGGTCTTACCCCATTTCA 1080
1081 TGGAGGGGTGAGTGAGGAGACAGATATATGATGCGCAGCATAAACAAATACACAA 1140
1081 TGGAGGGGTGAGTGAGGAGACAGATATATGATGCGCAGCATAAACAAATACACAA 1140
1141 CCCTAATTAACACTTCCCTCTTCTACTGACACCCCTTCACTCTCTCTTTCAATAAAAA 1200
1141 CCCTAATTAACACTTCCCTCTTCTACTGACACCCCTTCACTCTCTCTTTCAATAAAAA 1200
1201 TAAAAAAGATTTTATGTGGCTCTTACGATAGAAATCTTTCTCGAACTATAAAAAAGATC 1260
1201 TAAAAAAGATTTTATGTGGCTCTTACGATAGAAATCTTTCTCGAACTATAAAAAAGATC 1260

1261 TAAATATTTATATTTTCACTTTTAAATATCTTAGCGATGACAAAGCAGAAACAAATATT 1320
1261 TAAATATTTATATTTTCACTTTTAAATATCTTAGCGATGACAAAGCAGAAACAAATATT 1320
1321 TTTTGGCTCTCTCAACAGCAAAAGCTTGGGCTCTTTTGGTTCGGTGTAGGAATAGAA 1380
1321 TTTTGGCTCTCTCAACAGCAAAAGCTTGGGCTCTTTTGGTTCGGTGTAGGAATAGAA 1380
1381 CGAGAGCCCGGTATCTATAGGCGAGATGCTCTATCATATTAGCCCATGATCTCCAGCCTCAG 1440
1381 CGAGAGCCCGGTATCTATAGGCGAGATGCTCTATCATATTAGCCCATGATCTCCAGCCTCAG 1440
1441 AGCACAATTTTCTCGGGCTCTCTTAAAGCTTTTCCACAGCATTTGGGAACTTTTACTGAC 1500
1441 AGCACAATTTTCTCGGGCTCTCTTAAAGCTTTTCCACAGCATTTGGGAACTTTTACTGAC 1500
1501 AGCATCCAAAGTTGTCTTCTGTGAAGAACTGGAATCACAATCTCTGTGTCATCACTTCGG 1560
1501 AGCATCCAAAGTTGTCTTCTGTGAAGAACTGGAATCACAATCTCTGTGTCATCACTTCGG 1560
1561 CCCGTTTGGGTAGATCTCTGATTTAGGCTTTAGGCTTTAGAACACGGTGAGCCTGTGGT 1620
1561 CCCGTTTGGGTAGATCTCTGATTTAGGCTTTAGGCTTTAGAACACGGTGAGCCTGTGGT 1620
1621 GCACATAATTTATGCGGAGTGAACACCATAGAGTCAAGTGCATTTACTGAATGCTTCAATTT 1680
1621 GCACATAATTTATGCGGAGTGAACACCATAGAGTCAAGTGCATTTACTGAATGCTTCAATTT 1680
1681 CTCTTAATCTGATCGATGGCATGTACAGGGCCATTTTGTGTCAGACATCACTCCAG 1740
1681 CTCTTAATCTGATCGATGGCATGTACAGGGCCATTTTGTGTCAGACATCACTCCAG 1740
1741 AGAATTCAAAACAGATAGAGACAGTGGGACCCAGACCCATCTCTTCCCTCGGCTGA 1800
1741 AGAATTCAAAACAGATAGAGACAGTGGGACCCAGACCCATCTCTTCCCTCGGCTGA 1800
1801 TTATCCCGAGAAATAGGATGTCCCAAAGCAACACTTCCAGGCCAATGGAGTGTCTGATA 1860
1801 TTATCCCGAGAAATAGGATGTCCCAAAGCAACACTTCCAGGCCAATGGAGTGTCTGATA 1860
1861 GTCCAGTTATCAGAAAGATATGGCTGTAAAGTGTGATGCAAGTGTTCATTTTCTGAT 1920
1861 GTCCAGTTATCAGAAAGATATGGCTGTAAAGTGTGATGCAAGTGTTCATTTTCTGAT 1920
1921 AGTTAGTCTATAGAGCTGACAAAGAGGAAAGAGACGGGATGTGGTCAATTTA 1980
1921 AGTTAGTCTATAGAGCTGACAAAGAGGAAAGAGACGGGATGTGGTCAATTTA 1980
1981 ACAGGCGAGCTGTCCCTGGCTTCCCGATACGTGGGATGACTCGCATTTGTGAGCGGTG 2040
1981 ACAGGCGAGCTGTCCCTGGCTTCCCGATACGTGGGATGACTCGCATTTGTGAGCGGTG 2040
2041 GTCACCTGCAAAAGGAATGACCTCTCACATTTCTTCTGATTCGCATAGCGCGGCCAG 2100
2041 GTCACCTGCAAAAGGAATGACCTCTCACATTTCTTCTGATTCGCATAGCGCGGCCAG 2100
2101 CTTGTCTCTCTCTTGGGCTTCCAGACACTTAAGTCTGGAATGAAATTCACCTGCCT 2160
2101 CTTGTCTCTCTCTTGGGCTTCCAGACACTTAAGTCTGGAATGAAATTCACCTGCCT 2160
2161 CTGAATTTGGCCACTGTTGGGGCGAGGGGTGATCTTGGCTTCCAGGCTGGAAGATATC 2220
2161 CTGAATTTGGCCACTGTTGGGGCGAGGGGTGATCTTGGCTTCCAGGCTGGAAGATATC 2220
2221 TCACCCAGCCCTAGCTATATAACGGGCTGGTGGAGGGGCTTCCACAGGGCCAGTTCCAG 2280
2221 TCACCCAGCCCTAGCTATATAACGGGCTGGTGGAGGGGCTTCCACAGGGCCAGTTCCAG 2280
2281 GGGTTTCATCCACAGAGAGAAAAACATAGACTCGAGGTCTAGGAGCTTGCATGCTGCA 2340
2281 GGGTTTCATCCACAGAGAGAAAAACATAGACTCGAGGTCTAGGAGCTTGCATGCTGCA 2340
2341 GGTGCGAGGCCACCATGG 2358

```
Db 2341 GGTGGAGGCCACCATGG 2358
|||||
RESULT 2
US-10-005-337A-2
; Sequence 2, Application US/10005337A
; Publication No. US20030039984A1
; GENERAL INFORMATION:
; APPLICANT: BENOIT, Patrick
; APPLICANT: SCHWARTZ, Bertrand
; APPLICANT: BRANELLEC, Didier
; APPLICANT: CHIEN, Kenneth R.
; TITLE OF INVENTION: SEQUENCES UPSTREAM OF THE CARP GENE, VECTORS CONTAINING
; FILE REFERENCE: 03806.0530-00000
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US/10/005,337A
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2074
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-005-337A-2

Query Match 15.6%; Score 367.8; DB 14; Length 2074;
Best Local Similarity 76.4%; Pred. No. 3.4e-101; Indels 22; Gaps 11;
Matches 597; Conservative 0; Mismatches 162;

QY 1540 TCTCTCTGTGCATCAGTTCGGCCCGTTTGGGGT--AGATCCTCTGTAGTCCCTTCAGAT 1597
|||
Db 1289 TCTATTTCTTGACCACTCTGATCCATTTTGAAGTAAATGCCTCCATTTATGCTGT 1348
|||
QY 1598 TTGAACACGGTGAAGCTGTGGTGCATTAATATGCGCAGTGACACCATAGAGTCAAAGT 1657
|||
Db 1349 TTGAACACGGTGAAGCATGTGTGCTGTA---ATGGCCAGTGACATCATAAAGAAAAGT 1405
|||
QY 1658 GCATTACTGAATGCTTTCAATTTCTCTTAATGCTGTGACGATGGCATGTACAGGGCCAT 1717
|||
Db 1406 GCATTACTGAATGCTTTCAATGCTTTAATGATGGAAGTGCGATGTCTATGGGGCCTA 1465
|||
QY 1718 TTTAGTGTGACATCACTCCAGAGATTCCTCAACAGATAGAGACAGTGGCCACCCAGAC 1777
|||
Db 1466 TTTAGC--CGACATCACTCCAAGATTCCTCAACAGATAGACAGTGGCTTTAGGGC 1524
|||
QY 1778 CCATCTCTCCCTCGGGCTGATTTATCCCAAGAAATAGGATGTCCCAAGCAACACTTC 1837
|||
Db 1525 CCAGATCCCTTCCCTCAGGCTGTTTACCCAGGGAATAGGATGTCTGGGACAAAGTTTC 1584
|||
QY 1838 CCAGCCAACTGGATGCTGATAGTCCAGTTATCAGAAAGATATGGCTGTAAAGTGTGATG 1897
|||
Db 1585 C--CTAAGTGAAGTGTGTATAGTCTGCTTATCAGAAAGATATTAAGTGGGGTGTGATA 1641
|||
QY 1898 CACAGTGC--TTGCATTTCTTGATAGTTAGTCAATATGAGCTGACAAAGAGGAAAA 1955
|||
Db 1642 TGTAGGCACTACATTTCTTGATA--GGTAGTCAATGAAGCTGCAAAAGAA--AAAA 1698
|||
QY 1956 AGAGCAGCGATGGTGCATATTAACAGGCAGCTGTCCCTGGCTTCCCGATACGTGGG 2015
|||
Db 1699 AGGCAGTGTGTTGCAATGTCAACAGACAGCTGTCCCTGAC--TCTTGACAAATAGG 1757
|||
QY 2016 ATGACTCGCAATGTGAGCGGTGTGGTCACTGCCAAAGGAATGACCTCTCAATTTCTT 2075
|||
Db 1758 ATGACTTGTGATTTGCTGAGCGATGTGATCAACCAAGGAATGGCCCTCTCAATTTCTT 1817
|||
QY 2076 CCTGATTCGCATACGGCGGG-----CCAGCTTGTGATCTCCCTCTTGGGCTTCCGAC 2130
|||
Db 1818 CCTGATTCATATTCAGAGGGGTAGCTTGTCTCCCTCCCTCTTCTTTCAGCTTCCGAC 1877
|||
QY 2131 ACTAAGTCTGGAATGAAATTCACCTGCCTCTGAAATTGGCCACTGTGGTGGGGCGGGGTG 2190
|||
```

```
Db 1878 ACTAGTCTGGAATGAAATTCACCTGCCTCTGAGTTGGTCTCTAATGGGGCGGAGTG 1937.
|||||
QY 2191 TGACTTGGCTTCCAGGCTGGAAGATTATCTCACCAGCCCTAGCTATATAA--CGGGCTG 2249
|||||
Db 1938 TTACTTCGGTTCAGAGTTCGAAGATTATCTCACCAGCCCGCCAGCTATATAGCTGACCG 1997
|||||
QY 2250 GTGTGGAGGGCTCCACAGGGCCAGTTCAGGGGTTCATCCAAAGAGAGAAAAACATAG 2309
|||||
Db 1998 GTGTGGAGGGGGCCAGCAGGGCCAACTCCAGGGATTCTTTC--CACGACAGAAAAACATAC 2056
|||||
QY 2310 A 2310
Db 2057 A 2057

RESULT 3
US-10-175-523-80
; Sequence 80, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/10795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 80
; LENGTH: 90650
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-175-523-80

Query Match 4.3%; Score 100.4; DB 14; Length 90650;
Best Local Similarity 74.7%; Pred. No. 5.7e-18;
Matches 139; Conservative 0; Mismatches 46; Indels 1; Gaps 1;

QY 198 TCTAGGCGCAGCTTCTCAACCTGGGGCTCCAGCCCTTTGGGGGAATCAACAGCCCT 257
|||||
Db 56512 TCTATAGCAGTGGTCTCAACAGCTGGCTATCAACCCCTATGTGG--CTGAATGACCCCTT 56570
|||||
QY 258 TTACAGGGGTACATATCATCTATATCTATATGTCAAGTATTTTACATTTACGATTCGTAACA 317
|||||
Db 56571 TCATAGAGTCAATATCTGATATCTGATATCATATTTACATTTACATTTACATTTACATTTAC 56630
|||||
QY 318 GTAGCAAAATTCAGGTATGAAATAGCAATGAAATGAAATTTTATGATGAAAGGTCAACACA 377
|||||
Db 56631 GTAGGAAAGTTACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 56690
|||||
QY 378 ACATGA 383
Db 56691 TGAGGA 56696
|||||
```

```
RESULT 4
US-09-771-208-20
; Sequence 20, Application US/09771208
; Patent No. US20020155564A1
; GENERAL INFORMATION:
; APPLICANT: MEDRANO, JUAN
; APPLICANT: BRADFORD, ERIC
; APPLICANT: HORVAT, SIMON
; TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
; FILE REFERENCE: 407T-923710US
; CURRENT APPLICATION NUMBER: US/09/771,208
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 08/999,477
; PRIOR FILING DATE: 1997-12-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 659158
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (123459)..(123478)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (602466)..(602485)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (546398)..(547017)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (494715)..(494814)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (390986)..(391005)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (346860)..(346823)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (317174)..(317193)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (280353)..(280373)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (271829)..(271848)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (183872)..(183891)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (170625)..(170645)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (132680)..(132700)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; OTHER INFORMATION: n is a, c, g, or t
US-09-771-208-20
Query Match 3.8%; Score 88.8; DB 10; Length 659158;
Best Local Similarity 74.7%; Pred. No. 8.5e-14;
Matches 139; Conservative 0; Mismatches 42; Indels 5; Gaps 2;

QY 204 GCAGCAGTCTCAACCTGGGGCTCGACCCCTTTGGGGAATCAACGACCCCTTTACAG 263
Db 208828 GCAGCAGTCTCCNACGTGGGTCCAGATCTCTTTGGGAG-CCGAAAGGGCCCTTTCACAG 208886
QY 264 GGGTCACATCATCTCTATATGTCAGGTATTTACATTTAGCATTTGTAACAGTAGCA 323
Db 208887 GGGTCACAGTCAGATACCCCTGAATGTTGCATATTTATATTACAAATTCATAGCAGCTGCA 208946

RESULT 5
US-10-175-523-97
; Sequence 97, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/LJ795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 106664
; TYPE: DNA
; ORGANISM: Mus musculus domesticus
US-10-175-523-97
Query Match 3.6%; Score 84.2; DB 14; Length 106664;
Best Local Similarity 68.8%; Pred. No. 6e-13;
Matches 130; Conservative 0; Mismatches 58; Indels 1; Gaps 1;

QY 198 TCTAGGGCAGCAGTCTCAACCTGGGGCTCGACCCCTTTGGGGAATCAACGACCCCT 257
Db 35071 TCTAATACAAATGATTTCTTAATCTGTGGGAGTGGCCCTTTTAGGAGGTTCAAGGACCCCT 35130
QY 258 TTACAGGGGTCAATATCATCTATATGTCAGGTATTTACATTTAGCATTTACATTTAGTTAACA 317
Db 35131 TTCATAGGGGTTCCTAAGACCATCGGAAACATAGATATTTACATTTAGTTAATA 35190
QY 318 GTAGCAAAATACAGTATGAAATAGCAAT-GAAATAATTTATGATTTAGTGAAGTCCACC 376
Db 35191 GTAGCAAAATACATTTATGGAGTAACAATAAAATAATTTATGTTGGGGTCCACC 35250
QY 377 AACATGAGG 385
Db 35251 AACATAATG 35259

RESULT 6
US-09-771-208-20/c
; Sequence 20, Application US/09771208
; Patent No. US20020155564A1
; GENERAL INFORMATION:
; APPLICANT: MEDRANO, JUAN
; APPLICANT: BRADFORD, ERIC
```


QY 197 TTCTAGGGCAGCAGTTCTCAACCTGGGGCCCTCGACCCCTTTGGGGGAATCAACGACCC 256
DB 10340 TCGGAGGACAGAGTTCTCAACCTGTGGGGCACAACCCCTTTGGGG--TTTGAAGACTT 10397
QY 257 TTTACAGGGGTACATATCA---TCTATCTATATGTCTAGGTATTTACATTACGATTCT 313
DB 10398 TTTACAGGGGTTCCTTGCCCAAGACCATCAGAAAACACAGGTATTTACATTCCAATTCA 10457
QY 314 AACAGTAGCAAAATTACAGGTATGAATAGCAATGAATAATTTATGATTGAAGTCA 373
DB 10458 GACAGTAGTAAATTTACAGTTATGCAAGTATGCAATGAAATAATTTATGTTGGGTGAC 10517
QY 374 CACAACATGAGG 385
DB 10518 CACATCATGAAG 10529

RESULT 9

US-10-175-523-62
; Sequence 62, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, David
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Pallfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/LJ795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
; LENGTH: 185548
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-175-523-62

Query Match 3.3%; Score 78.2; DB 14; Length 185548;
Best Local Similarity 77.2%; Pred. No. 6.2e-11;
Matches 95; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 263 GGGGTACATATCATCTCTATATGTCTAGGTATTTACATTACGATTGTAACAGTAGC 322
DB 55602 GAGGTACACATCAGATATCCTCCATCAATATTTACATTATATTAATTTCAACACTAGC 55661
QY 323 AAAATTACAGGTATGAATAGCAATGAATAATTTATGATTGAAGGTCAACCAATG 382
DB 55662 AAAATTAAAGTTACAAAGTAGCAACAAAATAATTTATGTTAGGGGCCATCACAATG 55721
QY 383 AGG 385
DB 55722 AGG 55724

RESULT 10

US-10-228-794-75

; Sequence 75, Application US/10228794
; Publication No. US20030027198A1
; GENERAL INFORMATION:
; APPLICANT: VANDERBILT UNIVERSITY
; 305 Kirkland Hall
; Nashville, TN 37240
; TITLE OF INVENTION: MAMMALIAN GENES INVOLVED IN VIRAL
; INFECTION
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Needle & Rosenberg, P.C.
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/10/228,794
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,209
; FILING DATE: 08-Mar-1999
; APPLICATION NUMBER: PCT/US97/06067
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Selby, Elizabeth
; REGISTRATION NUMBER: 38,298
; REFERENCE/DOCKET NUMBER: 22000.0061/P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404 688 0770
; TELEFAX: 404 688 9880
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 892 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 75:
US-10-228-794-75

Query Match 3.3%; Score 78; DB 14; Length 892;
Best Local Similarity 70.5%; Pred. No. 1.5e-12;
Matches 148; Conservative 0; Mismatches 55; Indels 7; Gaps 3;
QY 177 CTCTTTTCAGGATCAGCTGATTCCTAGGGCAGCAGTTCTCAACCTGGGGCCCTCGACCCCT 236
DB 119 CAGGTACATAGATAGTCAAAATCTAGAGCACTGTTTCTATACCTGTGAGTTGCAACCCCT 178
QY 237 TTGGG---GGAATCAACGACCCCT--TTACAGGGGTACATATCATCTATCTATATGTCA 292
DB 179 TTGGGAGTGGGTCAAAATGACCCCTATCAGGGGTTCAAAATGAGATATCTTCATATCA 238
QY 293 GGTATTTACATTACGATTCGTAAACAGTAGCAAAATTTACAGGTATGAATAGCAATGAAAT 352
DB 239 AATATTTACATTATGATTCATAGTAGTACCAGATTTACAGTTATGAAGTTACA--AAAT 295
QY 353 AATTTTATGATTGAAGGTCAACCAACATG 382
DB 296 AATTTTATAGCTGAGAGTCACCAACATG 325

RESULT 11

US-10-026-188-3/c
; Sequence 3, Application US/10026188
; Publication No. US20020164645A1
; GENERAL INFORMATION:

```
; APPLICANT: Zuker, Charles S.
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Assays for Taste Receptor Cell Specific
; TITLE OF INVENTION: Ion Channel
; FILE REFERENCE: 02307E-114910US
; CURRENT APPLICATION NUMBER: US/10/026,188
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 60/259,379
; PRIOR FILING DATE: 2000-12-29
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 249487
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: mouse genomic region containing ltrpes
US-10-026-188-3

Query Match      3.3%; Score 76.8; DB 13; Length 249487;
Best Local Similarity 70.2%; Pred. No. 2e-10;
Matches 146; Conservative 0; Mismatches 57; Indels 5; Gaps 3;

QY 179 CTTTCAGGATCAGCCTGATTCTAGGGCAGCAGTCTCAACCTGGGGCCTCGACCCCTTT 238
Db 238599 CTTACAGCCTTTCTAGCCTCTACAGCAGTGGTTCTCAACCTGTGGTTTGAGACCCCTTT 238540

QY 239 GGGGGAATCAAGACCCCTTACAGGGGTACATATCATCTATCTATGTGTCAGGTATT 298
Db 238539 --GGAAGTCATATGACCCCTTCAAGGGCTTGCTCAAGACCATC--AGAAAAATAGATATT 238484

QY 299 TACATTACGATTCGTAAACAGTAGCAAAATTTACAGGTATGAAATAGCAAT-GAAATAATTT 357
Db 238483 TACATTATGATTCATAGCAGTAGCAAAATTTATAGTTATGAGTAGCAATAAATAATTT 238424

QY 358 TATGATTGAAGTCAACCAACATGAGG 385
Db 238423 TGTGTTGGAGATCACCACCAATGAGG 238396

RESULT 12
US-09-728-445-91
; Sequence 91, Application US/09728445
; Patent No. US20020102543A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020102543A1el Mutated Mammalian Cells and
; FILE REFERENCE: LEX-0102-USA
; CURRENT APPLICATION NUMBER: US/09/728,445
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,358
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 891
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 91
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(335)
; OTHER INFORMATION: n = A,T,C or G
US-09-728-445-91

Query Match      3.2%; Score 75.8; DB 10; Length 335;
Best Local Similarity 81.8%; Pred. No. 3.6e-12;
Matches 112; Conservative 0; Mismatches 22; Indels 3; Gaps 2;

QY 210 GTTCTCAACCTGGGGCCTCGACCCCTTTGGGGGAATCAACGACCCCTTT-ACAGGGGTC 268
```

```
Db 201 GTTCTCAACCTGGGGTTCACCCCTTTGAGGG--TCAACGACCCCTTTTCAAGAGTT 258
QY 269 ACATATCATCTATCTCTATATGTCCAGGTATTTCATTACGATTCTGTAACAGTAGCAAAATT 328
Db 259 ACGTATTAGCTATCTCTGGATGTCAGATATTTTACATTACATGTATATACAGCAGCAAAATT 318
QY 329 ACAGGTATGAATAGCA 345
Db 319 ACAGTTATGAAGTAGCA 335

RESULT 13
US-09-917-800A-477/c
; Sequence 477, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castile, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 477
; LENGTH: 5990
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AB009636
US-09-917-800A-477

Query Match      3.2%; Score 75.6; DB 10; Length 5990;
Best Local Similarity 66.7%; Pred. No. 3.3e-11;
Matches 124; Conservative 0; Mismatches 59; Indels 3; Gaps 1;

QY 196 ATTCTAGGCAGCAGTTCTCAACCTGGGGCCTCGACCCCTTTGGGGGAATCAACGACC 255
Db 4978 ACTCTAAGCCTGTGGTTCTCAACCCATGAGTCATGACCCCTTTGGGATTGTCCAAATGACCC 4919

QY 256 CTTT---ACAGGGTCACATATCATCTATCTATCTAGGTATTTCATTACGATTTCG 312
Db 4918 TTTTATCACAGGGGTGCAATAGTAGGTATCTCTGTAGATCAGATATATATATTCATTC 4859

QY 313 TAACAGTAGCAAAATTTACAGGTATGAATAGCAATGAAATAATTTTATGATTGATCAAGTCA 372
Db 4858 TAACCGTGTCAAAATTTACAGTTGTGTAATAGCACTAAATAATATTTTGTGTTGGTCA 4799

QY 373 CCACAA 378
```


Db 4798 CTACAA 4793

RESULT 14

US-09-901-484A-183

Sequence 183, Application US/09901484A

Patent No. US20020119460A1

GENERAL INFORMATION:

APPLICANT: Cohen, Daniel

APPLICANT: Blumenfeld, Marta

APPLICANT: Chumakov, Ilya

APPLICANT: Bouqueleret, Lydie

TITLE OF INVENTION: Prostate Cancer Gene

FILE REFERENCE: GEN-T111XG3D2

CURRENT APPLICATION NUMBER: US/09/901,484A

PRIOR FILING DATE: 2001-07-09

PRIOR APPLICATION NUMBER: US 08/996,306

PRIOR FILING DATE: 1997-12-22

PRIOR APPLICATION NUMBER: US 60/099,658

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: US 09/218,207

PRIOR FILING DATE: 1998-12-22

PRIOR APPLICATION NUMBER: US 09/338,907

PRIOR FILING DATE: 1999-06-23

PRIOR APPLICATION NUMBER: US 09/853,526

PRIOR FILING DATE: 2001-05-11

NUMBER OF SEQ ID NOS: 578

SOFTWARE: Patent in version 3.1

SEQ ID NO 183

LENGTH: 37950

TYPE: DNA

ORGANISM: Mus musculus

FEATURE:

NAME/KEY: misc_feature

LOCATION: (616)..(616)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (1552)..(1552)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (2809)..(2809)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (2821)..(2824)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (2826)..(2826)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (2828)..(2829)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (2831)..(2833)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (2835)..(2835)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (5259)..(5328)

OTHER INFORMATION: exon 2

NAME/KEY: misc_feature

LOCATION: (6247)..(6247)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (8667)..(8667)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (8669)..(8669)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (8671)..(8679)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (8681)..(8685)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (8687)..(8687)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (8689)..(8693)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (8698)..(8698)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (9062)..(9062)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (12298)..(12298)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (12675)..(12791)

OTHER INFORMATION: exon 3

NAME/KEY: misc_feature

LOCATION: (14621)..(14710)

OTHER INFORMATION: exon 4

NAME/KEY: misc_feature

LOCATION: (17578)..(17578)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (17639)..(17639)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (17707)..(17707)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (17867)..(17867)

OTHER INFORMATION: n = a, c, g, or t.

NAME/KEY: misc_feature

LOCATION: (19822)..(19912)

OTHER INFORMATION: exon 5

NAME/KEY: misc_feature

LOCATION: (21789)..(21950)

OTHER INFORMATION: exon 6

NAME/KEY: misc_feature

LOCATION: (23387)..(23510)

OTHER INFORMATION: exon 7

NAME/KEY: misc_feature

LOCATION: (25520)..(26016)

OTHER INFORMATION: exon 8

NAME/KEY: misc_feature

LOCATION: (37931)..(37931)

OTHER INFORMATION: n = a, c, g, or t.

US-09-901-484A-183

Query Match 3.2%; Score 74.8; DB 10; Length 37950;

Best Local Similarity 80.0%; Pred. No. 2.2e-10;

Matches 88; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 276 ATCTATCTTATGTCAGTATTTACATTCGTAACAGTACGTAACAGTA 335

Db 9570 ACCTATCTCTGCTTATCAGATAGTATGTAATGTAACAGCAGCAAAATCAGGT 9629

QY 336 TGAATAGCAATGAAATATTTATGATTGAAGGTCACCAACATGAGG 385

Db 9630 CCAATATCAAAAATATTTATGTTGAGGTCACCATACGTGAGG 9679

RESULT 15

US-09-853-526-183

Sequence 193, Application US/09853526

Patent No. US20020165345A1

GENERAL INFORMATION:

APPLICANT: Cohen, Daniel

APPLICANT: Blumenfeld, Marta

APPLICANT: Ilya, Chumakov


```
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/853,526
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 09/218,207
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 183
; LENGTH: 37950.
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5259..5328
; OTHER INFORMATION: exon2
; NAME/KEY: exon
; LOCATION: 12675..12791
; OTHER INFORMATION: exon3
; NAME/KEY: exon
; LOCATION: 14621..14710
; OTHER INFORMATION: exon4
; NAME/KEY: exon
; LOCATION: 19822..19912
; OTHER INFORMATION: exon5
; NAME/KEY: exon
; LOCATION: 21789..21950
; OTHER INFORMATION: exon6
; NAME/KEY: exon
; LOCATION: 23387..23510
; OTHER INFORMATION: exon7
; NAME/KEY: exon
; LOCATION: 25520..26016
; OTHER INFORMATION: exon8
; US-09-853-526-183
```

```
Query Match      3.2%; Score 74.8; DB 10; Length 37950;
Best Local Similarity 80.0%; Pred. No. 2.2e-10;
Matches 88; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY      276 ATCTATCCTATATGTCAGGTATTTACATTACGATTTCGTAAACAGTAGCAAAATTCAGGTA 335
      |||||
Db      9570 ACCTATCCTGCTTATCAGATAGTTACATTATGAATTTGAACAGCAGCAAAATCACAGTTA 9629

QY      336 TGAATAGCAATGAAATATTTTATGATTGAAGGTCACCACACATGAGG 385
      |||||
Db      9630 CGCAATATCAACAAATATTTTATGTTGAGGGTCACCATAACGTGAGG 9679
```

Search completed: November 14, 2003, 11:39:40
Job time : 703.036 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 14, 2003, 03:50:47 ; Search time 7536.5 Seconds
(without alignments)
11258.085 Million cell updates/sec

Title: US-10-005-337A-2
Perfect score: 2074
Sequence: 1 cgcagcaggttacttaatg.....acaagactcttcagccaac 2074

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_man.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a .

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2071.6	99.9	2074	6	AX468604	AX468604 Sequence
2	2071.6	99.9	2074	9	AF131884	AF131884 Homo sapi
3	1395	67.3	158357	9	AL365434	AL365434 Human DNA
4	1395	67.3	160350	2	AC074094	AC074094 Homo sapi
5	664.8	32.1	50111	9	AL590622	AL590622 Human DNA
6	367.8	17.7	2358	6	AX468603	AX468603 Sequence
7	367.8	17.7	142902	2	AC119234	AC119234 Mus muscu
8	365	17.6	229640	2	AC105469	AC105469 Rattus no
9	365	17.6	238344	2	AC097115	AC097115 Rattus no
10	340.8	16.4	723	10	AF478692	AF478692 Mus muscu
11	228	11.0	1901	6	AX322775	AX322775 Sequence
12	228	11.0	1901	6	BD094076	BD094076 Shear str
13	228	11.0	1901	9	HSRACINP	X83703 H.sapiens m
14	228	11.0	1901	11	G28603	G28603 human STS S
15	154.4	7.4	1940	4	AF131883	AF131883 Oryctolag
16	144.4	7.0	1889	6	AX281749	AX281749 Sequence
17	133.6	6.4	838	11	BV062893	BV062893 S212P6025
18	102.2	4.9	210651	2	AC026715	AC026715 Homo sapi
19	101	4.9	160901	2	AC120060	AC120060 Rattus no
20	101	4.9	235419	2	AC095979	AC095979 Rattus no
21	100.6	4.9	38476	9	U73024	U73024 Homo sapien
22	99	4.8	116803	2	AC097718	AC097718 Homo sapi
23	99	4.8	157665	9	AC104695	AC104695 Homo sapi
24	98.8	4.8	87402	9	HSJ437M21	AL049758 Human DNA
25	98.4	4.7	186870	9	AC026471	AC026471 Homo sapi
26	98.4	4.7	227074	2	AC122863	AC122863 Mus muscu
27	98.2	4.7	51529	9	AL591291	AL591291 Human DNA
28	98.2	4.7	148259	10	AC104099	AC104099 Mus muscu
29	98	4.7	172260	9	AC068946	AC068946 Homo sapi
30	97.8	4.7	152696	2	AC129985	AC129985 Homo sapi
31	97.8	4.7	153053	9	AC090711	AC090711 Homo sapi
32	97.8	4.7	163433	2	AC079993	AC079993 Homo sapi
33	97.8	4.7	179040	2	AC022732	AC022732 Homo sapi
34	97.6	4.7	51370	2	AC084305	AC084305 Homo sapi
35	97.6	4.7	218309	2	AC137376	AC137376 Rattus no
36	97.6	4.7	249682	2	AC096394	AC096394 Rattus no
37	97.4	4.7	5404	9	HSM804865	AL833552 Homo sapi
38	97.4	4.7	34917	9	AL390211	AL390211 Human DNA
39	97.2	4.7	34488	6	BD093713	BD093713 Human lp3
40	97.2	4.7	90289	2	AP003164	AP003164 Homo sapi
41	97.2	4.7	121162	6	AX039602	AX039602 Sequence
42	97.2	4.7	171612	9	AL358013	AL358013 Human DNA
43	97	4.7	124321	9	HS417G15	AL009174 Human DNA
44	97	4.7	163338	9	AL162426	AL162426 Human DNA
45	96.8	4.7	142952	2	AC040909	AC040909 Homo sapi

ALIGNMENTS

RESULT 1
AX468604
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

AX468604
Sequence 2 from Patent WO0246220.
AX468604
AX468604.1 GI:21901403
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Schwartz,B., Branellec,D. and Chien,K.
Sequences upstream of the carp gene, vectors containing them and
uses thereof

AX468604
Sequence 2 from Patent WO0246220.
AX468604
AX468604.1 GI:21901403
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Schwartz,B., Branellec,D. and Chien,K.
Sequences upstream of the carp gene, vectors containing them and
uses thereof

AX468604
Sequence 2 from Patent WO0246220.
AX468604
AX468604.1 GI:21901403
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Schwartz,B., Branellec,D. and Chien,K.
Sequences upstream of the carp gene, vectors containing them and
uses thereof

AX468604
Sequence 2 from Patent WO0246220.
AX468604
AX468604.1 GI:21901403
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Schwartz,B., Branellec,D. and Chien,K.
Sequences upstream of the carp gene, vectors containing them and
uses thereof

JOURNAL Patent: WO 0246220-A 2 13-JUN-2002; Aventis Pharma S.A. (FR) ; The Regents of The University of California at San Diego (US) ; Benoit, Patrick (FR)									
FEATURES Location/Qualifiers									
1..2074									
/organism="Homo sapiens"									
/mol_type="genomic DNA"									
/db_xref="taxon:9606"									
BASE COUNT 612 a 469 c 416 g 572 t 5 others									
ORIGIN									
Query Match 99.9%; Score 2071.6; DB 6; Length 2074;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2074; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	CTGCGAGCAAGTTACTTAATGTTTTTGGCTCGACGATCCTCTCTGTAAATAATGAGAGATTAA	60						
Db	1	CTGCGCAAGTTACTTAATGTTTTTGGCTCGACGATCCTCTCTGTAAATAATGAGAGATTAA	60						
QY	61	GTCTTGTCTCCAACTTCGAGGGCATGGACAGCTCTGGGATTTTCATATCCAAGACCCCTTAAA	120						
Db	61	GTCTTGTCTCCAACTTCGAGGGCATGGACAGCTCTGGGATTTTCATATCCAAGACCCCTTAAA	120						
QY	121	CATCCACAGTCTCTCCGCCAAACACTTCTCCTCTCTAATACCTCCCTCAGTTTGGGTGAG	180						
Db	121	CATCCACAGTCTCTCCGCCAAACACTTCTCCTCTCTAATACCTCCCTCAGTTTGGGTGAG	180						
QY	181	GCCTGGAAACAAAAGGCATACGAAATGGTAGAAAAAGTGCATGACTACTCTCTGACTTTA	240						
Db	181	GCCTGGAAACAAAAGGCATACGAAATGGTAGAAAAAGTGCATGACTACTCTCTGACTTTA	240						
QY	241	GATGAGAGACCAATGAATAATAGTAGTCTCTGCTCTAATACCTCCCTCAGTTTGGGTGAG	300						
Db	241	GATGAGAGACCAATGAATAATAGTAGTCTCTGCTCTAATACCTCCCTCAGTTTGGGTGAG	300						
QY	301	ATAGAGAGCTATACAAAGAGATTAGCATGCACTCTGTGCAAGAATGACACACAAATTTGT	360						
Db	301	ATAGAGAGCTATACAAAGAGATTAGCATGCACTCTGTGCAAGAATGACACACAAATTTGT	360						
QY	361	GAACAATTCATATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	420						
Db	361	GAACAATTCATATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	420						
QY	421	AAAATAGTATAGCTGTGTCATCTCAAGAAAGCCAGAGATTTCCTTTATTATACCC	480						
Db	421	AAAATAGTATAGCTGTGTCATCTCAAGAAAGCCAGAGATTTCCTTTATTATACCC	480						
QY	481	CCTTTAAGATAGATAATTAGGAGACCGGAACATATGATACAGAGGTACTGGGAGGTCC	540						
Db	481	CCTTTAAGATAGATAATTAGGAGACCGGAACATATGATACAGAGGTACTGGGAGGTCC	540						
QY	541	CTCTTTGTCAATGTTTTGTCTTTGGGTGGGAGTGGATGTCCTCTCAAAAGTTTCAGAAAC	600						
Db	541	CTCTTTGTCAATGTTTTGTCTTTGGGTGGGAGTGGATGTCCTCTCAAAAGTTTCAGAAAC	600						
QY	601	ACATCCACTGACTGAGCATTTCAAGGGCAAGAGAGAAATGGCAGGCACATTGTTGATT	660						
Db	601	ACATCCACTGACTGAGCATTTCAAGGGCAAGAGAGAAATGGCAGGCACATTGTTGATT	660						
QY	661	GGGTGAGTTTGGGAGAAATAGACACACAAAGGTCAAAATCTTCTTAATTTAACACTT	720						
Db	661	GGGTGAGTTTGGGAGAAATAGACACACAAAGGTCAAAATCTTCTTAATTTAACACTT	720						
QY	721	CCCTCATTTACAAATTTCCCTTCTTCCCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	780						
Db	721	CCCTCATTTACAAATTTCCCTTCTTCCCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	780						
QY	781	AGTTTTTCTGAAACTATATAAATAATACCCAGTAGTGTTTACATAATTTACCTCAAGA	840						
Db	781	AGTTTTTCTGAAACTATATAAATAATACCCAGTAGTGTTTACATAATTTACCTCAAGA	840						
QY	841	TTAGAAACCAAGATAGACCTTTTCAACCTTCCGGAAGCAAGTGCATTTATCTCCTCC	900						
Db									

Db	841	TTAGAAACCAAGATAGACCTTTTCAACCTTCCGGAAGCAAGTGCATTTATCCTCC	900						
QY	901	AGCCACGCTCTCAAACTTGTATGATCATCAGAAATCATCTGGGTGCTTTTAAATTTCAAGATG	960						
Db	901	AGCCACGCTCTCAAACTTGTATGATCATCAGAAATCATCTGGGTGCTTTTAAATTTCAAGATG	960						
QY	961	ATTCTACGAGTTACATAAAATCAACTCAGAAATTCCTCGAGTGGGGCCAGGATCTGTA	1020						
Db	961	ATTCTACGAGTTACATAAAATCAACTCAGAAATTCCTCGAGTGGGGCCAGGATCTGTA	1020						
QY	1021	TTTCTGACAAAGCTCCCA CAGGTGATTCCTTTCCCCACAGCATTTTGAACCTTCAAGTCAA	1080						
Db	1021	TTTCTGACAAAGCTCCCA CAGGTGATTCCTTTCCCCACAGCATTTTGAACCTTCAAGTCAA	1080						
QY	1081	TGACCTAATCAGAGTCCTGCCATTCCTAATATCTGGTCTCATTTTBTCTATATATATA	1140						
Db	1081	TGACCTAATCAGAGTCCTGCCATTCCTAATATCTGGTCTCATTTTBTCTATATATATA	1140						
QY	1141	TAGTATTTTGTGTAGAGATGGGATTTTGCATTTTCCACAGCTAGTATTTGAACCTCTAA	1200						
Db	1141	TAGTATTTTGTGTAGAGATGGGATTTTGCATTTTCCACAGCTAGTATTTGAACCTCTAA	1200						
QY	1201	GCTAAGCAATCTTCTGCTCTCTGCTCCCAAAATGTTGGGATTA CAGGTGTAAGCCACTG	1260						
Db	1201	GCTAAGCAATCTTCTGCTCTCTGCTCCCAAAATGTTGGGATTA CAGGTGTAAGCCACTG	1260						
QY	1261	CACCCGGCTGATAGTGGTTCCTTACTTCTTCTTGTGACCACTCTGATCCATTTTGA	1320						
Db	1261	CACCCGGCTGATAGTGGTTCCTTACTTCTTCTTGTGACCACTCTGATCCATTTTGA	1320						
QY	1321	GTAAAAATGCTCCAAATTTATGCTGTTTTAGAAACGGTAAGCATGT CATGTGCTAAATG	1380						
Db	1321	GTAAAAATGCTCCCAATTTATGCTGTTTTAGAAACGGTAAGCATGT CATGTGCTAAATG	1380						
QY	1381	GCCAGTGATCATATAAAGAAAGTGCATTTACTGAATGCTTTCAATGTCTTATAATGATG	1440						
Db	1381	GCCAGTGATCATATAAAGAAAGTGCATTTACTGAATGCTTTCAATGTCTTATAATGATG	1440						
QY	1441	GTAAAGTGGCATGTCTATGGGGCTATTAGCCGACATCATCTCCAAAGAAATTCCAAAAC	1500						
Db	1441	GTAAAGTGGCATGTCTATGGGGCTATTAGCCGACATCATCTCCAAAGAAATTCCAAAAC	1500						
QY	1501	GATATAGCAAGTGGCTTTAGGGCCAGATCCCTTCCCTCAGGCTGTTTACCCAGGGAA	1560						
Db	1501	GATATAGCAAGTGGCTTTAGGGCCAGATCCCTTCCCTCAGGCTGTTTACCCAGGGAA	1560						
QY	1561	TAGGATGTCTGGGACAAAGTTTCCCTTAAGTGAAGTGTGATAAGTCTGCTTATCAGAAA	1620						
Db	1561	TAGGATGTCTGGGACAAAGTTTCCCTTAAGTGAAGTGTGATAAGTCTGCTTATCAGAAA	1620						
QY	1621	GATATTTAGTGGGGTGTGATGTAGGGCATCTACATTTTCTTGTATAGTGTATATGA	1680						
Db	1621	GATATTTAGTGGGGTGTGATGTAGGGCATCTACATTTTCTTGTATAGTGTATATGA	1680						
QY	1681	AGCTGTGACAAAGAAAGAAAGGCGAGTGTGCGTCAATGTCAACAGACAGCTGTCCCCT	1740						
Db	1681	AGCTGTGACAAAGAAAGAAAGGCGAGTGTGCGTCAATGTCAACAGACAGCTGTCCCCT	1740						
QY	1741	GACTCTTGACAAATAGGATGACTTGCATTTGCTGAGCGATGTGATCACCAAGAAATG	1800						
Db	1741	GACTCTTGACAAATAGGATGACTTGCATTTGCTGAGCGATGTGATCACCAAGAAATG	1800						
QY	1801	GCCCTCTCACATTTCTTCTGATTTACATATTTACAGAGGGTTAGCTTGTCTCCCTCCC	1860						
Db	1801	GCCCTCTCACATTTCTTCTGATTTACATATTTACAGAGGGTTAGCTTGTCTCCCTCCC	1860						
QY	1861	TCTTCAGCTTCCAGACACTGAGTCTGGAATGGAATTTACCTGTGCTGTGAGTTGCTCC	1920						
Db	1861	TCTTCAGCTTCCAGACACTGAGTCTGGAATGGAATTTACCTGTGCTGTGAGTTGCTCC	1920						
QY	1921	TAATGGGGGGGAGTGTGTTTACTTTCGGTTCCAGGTTGGAAGATTATCTCACCGGGCCCCA	1980						
Db	1921	TAATGGGGGGGAGTGTGTTTACTTTCGGTTCCAGGTTGGAAGATTATCTCACCGGGCCCCA	1980						

QY 1981 GCTATATAGCTGACCGGTGTGAGGGCCAGAGGCCCACTCCAGGGATTCCCTTCCA 2040
 DB 1981 GCTATATAGCTGACCGGTGTGAGGGCCAGAGGCCCACTCCAGGGATTCCCTTCCA 2040

QY 2041 CGACAGAAAACATACAGAGCTCTTCAGGCCAAC 2074
 DB 2041 CGACAGAAAACATACAGAGCTCTTCAGGCCAAC 2074

RESULT 2
 AF131884 2074 bp DNA linear PRI 09-FEB-2000
 LOCUS Homo sapiens cardiovascular-specific cardiac ankyrin repeat protein
 DEFINITION (CVARP) gene, 5'-flanking region and exon 1.
 ACCESSION AF131884
 VERSION AF131884.1 GI:6940841
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2074)
 AUTHORS Aihara,Y., Kurabayashi,M., Tanaka,T., Sekiguchi,K., Tomaru,K.,
 Kanai,H., Takeda,S. and Nagai,R.
 TITLE Human CVARP 5'-flanking region
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2074)
 AUTHORS Aihara,Y.
 TITLE Direct Submission
 JOURNAL Submitted (26-FEB-1999) Second Department of Internal Medicine,
 Gunma University School of Medicine, 3-39-15 Showa-machi, Maebashi
 371-8511, Japan

FEATURES
 source 1..2074
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 misc_feature 1..1832
 /notes="5'-flanking region"
 gene 1833..>2074
 /gene="CVARP"
 mRNA 1833..>2074
 /gene="CVARP"
 /product="cardiovascular-specific cardiac ankyrin repeat
 protein"
 exon /note="CARP"
 1833..2074
 /gene="CVARP"
 /number=1

BASE COUNT 612 a 469 c 416 g 572 t 5 others
 ORIGIN

Query Match 99.9%; Score 2071.6; DB 9; Length 2074;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2074; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGAGCAAGTTACTTAATGTTTTTGGCTCAGCATCCTCTGTAAATGAGAGCATTA 60
 DB 1 CTGAGCAAGTTACTTAATGTTTTTGGCTCAGCATCCTCTGTAAATGAGAGCATTA 60

QY 61 GTCTTGCTCCAACTTCGAGGGCATGGACGCTCTGGGATTTTCATATCCAGAGCCCTTAA 120
 DB 61 GTCTTGCTCCAACTTCGAGGGCATGGACGCTCTGGGATTTTCATATCCAGAGCCCTTAA 120

QY 121 CATCCACAGTCCTTCCCCCAACACATTCCTCCTAAATACCTCCCTCAGTTGGGTGAG 180
 DB 121 CATCCACAGTCCTTCCCCCAACACATTCCTCCTAAATACCTCCCTCAGTTGGGTGAG 180

QY 181 GCCTGGAAACAAAAGGCATACGAAATGGTAGAAAAAGTGCCATGACTCTTCTGACTTA 240
 DB 181 GCCTGGAAACAAAAGGCATACGAAATGGTAGAAAAAGTGCCATGACTCTTCTGACTTA 240

QY 241 GATGAAGAGACCAATGAAATAGTACTGTGTTTCTTTCAGCAGGACATATACTAAA 300
 DB 241 GATGAAGAGACCAATGAAATAGTACTGTGTTTCTTTCAGCAGGACATATACTAAA 300

QY 301 ATAGGAGCTATACAAAGAGATTAGCATGAGCTCTGTGCAAGAAATGACACACAAATTTGT 360
 DB 301 ATAGGAGCTATACAAAGAGATTAGCATGAGCTCTGTGCAAGAAATGACACACAAATTTGT 360

QY 361 GAAACATTCATATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 420
 DB 361 GAAACATTCATATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 420

QY 421 AAAATAGTACTGTGTGCTCATCTCAAGAAAGCCAGGAGATTTCCTTTATTTACCC 480
 DB 421 AAAATAGTACTGTGTGCTCATCTCAAGAAAGCCAGGAGATTTCCTTTATTTACCC 480

QY 481 CCTTTAAGATAGAAATATTAGGAGACCGGAACATATGATACAGGAGGTACTGGGAGGTCC 540
 DB 481 CCTTTAAGATAGAAATATTAGGAGACCGGAACATATGATACAGGAGGTACTGGGAGGTCC 540

QY 541 CTCTTTGTCAATGTTTTTGTCTTGGGGTGGGAGTCGATGCTCTCTCAAAAGTTTCAGAAAC 600
 DB 541 CTCTTTGTCAATGTTTTTGTCTTGGGGTGGGAGTCGATGCTCTCTCAAAAGTTTCAGAAAC 600

QY 601 ACCATCCACTGACTGAGCATTTCAAGGGCAAGAGGAGAAATGGCAGGCACATTTGTTGATT 660
 DB 601 ACCATCCACTGACTGAGCATTTCAAGGGCAAGAGGAGAAATGGCAGGCACATTTGTTGATT 660

QY 661 GGGTGAGTTTGGGAGAAATAGACACACAAAGGTCAACATACTCTTCTTAATTAACACTT 720
 DB 661 GGGTGAGTTTGGGAGAAATAGACACACAAAGGTCAACATACTCTTCTTAATTAACACTT 720

QY 721 CCTCTCATTCACAAATTCCTCTCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
 DB 721 CCTCTCATTCACAAATTCCTCTCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780

QY 781 AGTTTTTCTGAAACTATATAAAATACCCCGATGATGTTTACATAAATTTACACTCAAGA 840
 DB 781 AGTTTTTCTGAAACTATATAAAATACCCCGATGATGTTTACATAAATTTACACTCAAGA 840

QY 841 TTAGAAACCAAGAAATAGACACCTTTTCAACCTTCCGGAAGCAAGTGCATTTCCCTCC 900
 DB 841 TTAGAAACCAAGAAATAGACACCTTTTCAACCTTCCGGAAGCAAGTGCATTTCCCTCC 900

QY 901 AGCCAGTGTCTCAAAATCTTGATGATCAGATCATCTGGGTGCTTCTTCAAAATTTCAAGATG 960
 DB 901 AGCCAGTGTCTCAAAATCTTGATGATCAGATCATCTGGGTGCTTCTTCAAAATTTCAAGATG 960

QY 961 ATTCTTACGAGTTACCATAAATCAACTCAGAAATTCCTCGAGTGGGGCCAGGATCTGTA 1020
 DB 961 ATTCTTACGAGTTACCATAAATCAACTCAGAAATTCCTCGAGTGGGGCCAGGATCTGTA 1020

QY 1021 TTTCTGCAAGCTCCCAAGGTGATTCCTTTCCCAAGCATTTTGAAGACTTCAGCTCAA 1080
 DB 1021 TTTCTGCAAGCTCCCAAGGTGATTCCTTTCCCAAGCATTTTGAAGACTTCAGCTCAA 1080

QY 1081 TGACCTAATCAGAGTCTGCCATTTGCTAATATCTGCTCTCATTTTBTCTATATATATA 1140
 DB 1081 TGACCTAATCAGAGTCTGCCATTTGCTAATATCTGCTCTCATTTTBTCTATATATATA 1140

QY 1141 TAGTATTTGTGGTAGAGATGGGATTTTGGCATTTGCCAGGCTAGTATTGAACCTCTAA 1200
 DB 1141 TAGTATTTGTGGTAGAGATGGGATTTTGGCATTTGCCAGGCTAGTATTGAACCTCTAA 1200

QY 1201 GCTAAGCAATCTTCTGTCTGTGCTCCCAAAATGTTGGGATTAAGGTGAAGCCACTG 1260
 DB 1201 GCTAAGCAATCTTCTGTCTGTGCTCCCAAAATGTTGGGATTAAGGTGAAGCCACTG 1260

QY 1261 CACCCGGCTGATAGTGGTTTCAATTTACTCTATTTCTTGACCACTCTGATCCATTTTGA 1320
 DB 1261 CACCCGGCTGATAGTGGTTTCAATTTACTCTATTTCTTGACCACTCTGATCCATTTTGA 1320

QY 1321 GTAAAAATGCTCCAATTATTATGCTGTTTTAGAAACACGGTAAGCATGTCTATGCTAATG 1380

```

Db      1321 GTAAAAATGCTCAATTAATGCTGTTTAGAACACGCTAGCATGTCATGCTAATG 1380
QY      1381 GCCAGTGACATCAATAAGAAAGTGCATTAAGTGAATGTTCAATGTTTAAATGATG 1440
Db      1381 GCCAGTGACATCAATAAGAAAGTGCATTAAGTGAATGTTTCAATGTTTAAATGATG 1440
QY      1441 GTAAGGTGGCATGTCTATGGGGCTATTAGCCGACATCACTCCAAAGAAATCCAAACA 1500
Db      1441 GTAAGGTGGCATGTCTATGGGGCTATTAGCCGACATCACTCCAAAGAAATCCAAACA 1500
QY      1501 GATATAGACAAGTGCCTTTAGGGCCAGATCCCTTCCCTCAGGCTGTTTACCCAGGAA 1560
Db      1501 GATATAGACAAGTGCCTTTAGGGCCAGATCCCTTCCCTCAGGCTGTTTACCCAGGAA 1560
QY      1561 TAGGATGTCCTGGGCAAGTTCCTTAAGTGAAGTGTGTAAGTCTGCTTATCAGAAA 1620
Db      1561 TAGGATGTCCTGGGCAAGTTCCTTAAGTGAAGTGTGTAAGTCTGCTTATCAGAAA 1620
QY      1621 GATATTACTGGGGTGTATATGTAGGGCATCTACATTTCTTGATAGTAGTCAATGA 1680
Db      1621 GATATTACTGGGGTGTATATGTAGGGCATCTACATTTCTTGATAGTAGTCAATGA 1680
QY      1681 AAGCTGACAAAGAAAAAGGCGAGTGATGTGTCGAATGTCAACAGACAGCTGTCCCT 1740
Db      1681 AAGCTGACAAAGAAAAAGGCGAGTGATGTGTCGAATGTCAACAGACAGCTGTCCCT 1740
QY      1741 GACTCTTGACAAATAGGATGACTTGCTGAGCGATGTGATCACCACCAAGGAATG 1800
Db      1741 GACTCTTGACAAATAGGATGACTTGCTGAGCGATGTGATCACCACCAAGGAATG 1800
QY      1801 GCCTCTCACATTTCTTCCTGATTCATATATCAGCAGGTTAGTGTCTCCCTCC 1860
Db      1801 GCCTCTCACATTTCTTCCTGATTCATATATCAGCAGGTTAGTGTCTCCCTCC 1860
QY      1861 TCTTCAGCTTCCAGACACTGAGTCTGGAATGAAATTCACCTGCTCTGAGTTGGCTCC 1920
Db      1861 TCTTCAGCTTCCAGACACTGAGTCTGGAATGAAATTCACCTGCTCTGAGTTGGCTCC 1920
QY      1921 TAATGGGGGGAGTGTACTTTCGTTCCAGTTCCAGTTGGAAGATTTATCTACCCCGGCCA 1980
Db      1921 TAATGGGGGGAGTGTACTTTCGTTCCAGTTGGAAGATTTATCTACCCCGGCCA 1980
QY      1981 GCTATATAAGCTGACCGGTGTGAGGGGCCAGCAGGCGCACTCCAGGATTCCTTCA 2040
Db      1981 GCTATATAAGCTGACCGGTGTGAGGGGCCAGCAGGCGCACTCCAGGATTCCTTCA 2040
QY      2041 CGACAGAAAAACATACAGACTCCTTTCAGCCCAAC 2074
Db      2041 CGACAGAAAAACATACAGACTCCTTTCAGCCCAAC 2074

```

```

RESULT 3
AL365434/c
LOCUS   158357 bp      DNA      linear      PRI 21-JUN-2002
DEFINITION Human DNA sequence from clone RP11-236B18 on chromosome 10,
           complete sequence.
ACCESSION AL365434
VERSION   AL365434.13  GI:21540024
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 158357)
AUTHORS   Bird, C.
TITLE     Direct Submission
JOURNAL   Submitted (21-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
           Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
           humquery@sanger.ac.uk
           On Jun 23, 2002 this sequence version replaced gi:12191663.
           During sequence assembly data is compared from overlapping clones.
           Where differences are found these are annotated as variations

```

together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human

chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr10

RP11-236B18 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pBAC3.6

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest.

Location/Qualifiers

1..158357

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="10"

/clone="RP11-236B18"

/clone_lib="RP11-11.1"

/base_count 46958 a 32356 c 32839 g 46204 t

BASE COUNT

ORIGIN

Query Match 67.3%; Score 1395; DB 9; Length 158357;

Best Local Similarity 98.4%; Pred. No. 7,2e-307;

Matches 1479; Conservative 4; Mismatches 10; Indels 10; Gaps 7;

QY 1 CTGACGAAAGTTACTTAATGTTT-TTGGCTCAGCATCCTCTCTTAAATGAGAGCATT 59

Db 1502 CTGACGAAAGTTACTTAATGTTTCTTGGCTCAGCATCCTCTCTTAAATGAGAGCATT 1443

QY 60 AGTCTTGTCTCAACTTCGAGGGCATGGACAGCTCTGGGATTTTCATATCCAGACCTTAA 119

Db 1442 AGTCTTGTCTCAACTTCGAGGGCATGGACAGCTCTGGGATTTTCATATCCAGACCTTAA 1383

QY 120 ACATCCCAAGCTCTTCCCAACACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 179

Db 1382 ACATCCCAAGCTCTTCCCAACACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1323

QY 180 GGCCTGGAAACAAAAGGCATACGAAATGGTGAAGAAAGTCCATGACTTCTTCTGACTT 239

Db 1322 GGCCTGGAAACAAAAGGCATACGAAATGGTGAAGAAAGTCCATGACTTCTTCTGACTT 1263

QY 240 AGATGAAGACCAATGAAATAGTAACTGCTGTTTCTTCTCAGCAGACATATATACTAA 299

Db 1262 AGATGAAGACCAATGAAATAGTAACTGCTGTTTCTTCTCAGCAGACATATATACTAA 1203

QY 300 AATAGGAGCTATACAAAGAGATTAGCATGCTCTGTGCAAGAAATGACACAAATTTG 359

Db 1202 AATAGGAGCTATACAAAGAGATTAGCATGCTCTGTGCAAGAAATGACACAAATTTG 1143

QY 360 TGAACATTTCCATATATTTAAATAATAATAATAATAATAATAATAATAATAATAATAATA 419

Db 1142 TGAACATTTCCATATATTTAAATAATAATAATAATAATAATAATAATAATAATAATAATA 1083

* 113398 113497: gap of unknown length
 * 113498 141732: contig of 28235 bp in length
 * 141733 141832: gap of unknown length
 * 141833 143187: contig of 1355 bp in length
 * 143188 143287: gap of unknown length
 * 143288 144942: contig of 1655 bp in length
 * 144943 145042: gap of unknown length
 * 145043 147286: contig of 2244 bp in length
 * 147287 147386: gap of unknown length
 * 147387 150852: contig of 3466 bp in length
 * 150853 150952: gap of unknown length
 * 150953 153550: contig of 2598 bp in length
 * 153551 153650: gap of unknown length
 * 153651 157257: contig of 3607 bp in length
 * 157258 157357: gap of unknown length
 * 157358 160350: contig of 2993 bp in length.

FEATURES

source

1. 160350
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP11-236B18"
 misc_feature 1. 2326
 /note="assembly_name:Contig10"
 misc_feature 2427..7519
 /note="assembly_name:Contig11"
 misc_feature 7620..15286
 /note="assembly_name:Contig12"
 misc_feature 15387..23050
 /note="assembly_name:Contig13"
 misc_feature 23151..33663
 /note="assembly_name:Contig14"
 misc_feature 33764..50133
 /note="assembly_name:Contig15"
 misc_feature 50234..68088
 /note="assembly_name:Contig16"
 misc_feature 68189..88133
 /note="assembly_name:Contig17"
 misc_feature 88234..113397
 /note="assembly_name:Contig18"
 misc_feature 113498..141732
 /note="assembly_name:Contig19"
 misc_feature 141833..143187
 /note="assembly_name:Contig3"
 misc_feature 143288..144942
 /note="assembly_name:Contig4"
 misc_feature 145043..147286
 /note="assembly_name:Contig5"
 misc_feature 147387..150852
 /note="assembly_name:Contig6"
 misc_feature 150953..153550
 /note="assembly_name:Contig7"
 misc_feature 153651..157257
 /note="assembly_name:Contig8"
 misc_feature 157358..160350
 /note="assembly_name:Contig9"
 BASE COUNT 47390 a 32724 c 33135 g 45498 t 1603 others
 ORIGIN

Query Match 67.3%; Score 1395; DB 2; Length 160350;
 Best Local Similarity 92.0%; Pred. No. 7.2e-307;
 Matches 1482; Conservative 4; Mismatches 115; Indels 10; Gaps 7;
 QY 1 CTGCAGCAAGTTACTTAATGTTT-TTGGCTCAGCATCCTCTCTGTGTAATAATCAGAGCATT 59
 DB 114999 CTGCAGCAAGTTACTTAATGTTTCTTGGCTCAGCATCCTCTCTGTGTAATAATCAGAGCATT 114940
 QY 60 AGTCTTGCTCCAACTTCGAGGGCATGACAGCTCTGGGATTCATATCCAAAGCCCTTAA 119
 DB 114939 AGTCTTGCTCCAACTTCGAGGGCATGACAGCTCTGGGATTCATATCCAAAGCCCTTAA 114980
 QY 120 ACATCCACAGTCTCTCCCAACACACTTCTCCTCCTAATACTCCTCCTCAGTTTGGGTCA 179

Db 114879 ACATCCACAGTCTCTTCCCAACACACTTCTCCTAATACCTCCCTCAGTTTGGGTCA 114820
 QY 180 GGCCTGGACAAAGAGGATAGCAATGGTGTAGAAAAGTGTCCATGACTTCTTCTGACTT 239
 Db 114819 GGCCTGGACAAAGAGGATAGCAATGGTGTAGAAAAGTGTCCATGACTTCTTCTGACTT 114760
 QY 240 AGATGAAGAGACCAATGAAATAGTAATGACTCTCTGTTGCTTCCAGCAGGACATATACTAA 299
 Db 114759 AGATGAAGAGACCAATGAAATAGTAATGACTCTCTGTTGCTTCCAGCAGGACATATACTAA 114700
 QY 300 AATAGGAGCTATACAAAGAAGATTAGCATGGACTCTGTGCAAGAATGACACAAATTTG 359
 Db 114699 AATAGGAGCTATACAAAGAAGATTAGCATGGACTCTGTGCAAGAATGACACAAATTTG 114640
 QY 360 TCAAACTTCATATATTAATAATAATAATAATAAGAGAAAAGAAAATTAATAAA 419
 Db 114639 TGAACATTTCCATATATTAATAATAATAATAATAAGAGAAAAGAAAATTAATAAA 114580
 QY 420 GAAATAGTATAGTGTGTCCATCTCAAGAAAAGCCAGGAGATTCTCTTTATTTA-- 477
 Db 114579 GAAATAGTATAGTGTGTCCATCTCAAGAAAAGCCAGGAGATTCTCTTTATTTA 114520
 QY 478 CCCCTTTAAGATAGATAATATTAGGAGACCGGAACATATATACAGGAGGTACTGGAGGG 537
 Db 114519 CCCCTTTAAGATAGATAATATTAGGAGACCGGAACATATATACAGGAGGTACTGGAGGG 114460
 QY 538 TCCCTTTTGCAATGTTTGTCTTGGGTGGGAGTGCATGCTCTTCAAAAGTTTCAG 597
 Db 114459 TCCCTTTTGCAATGTTTGTCTTGGGTGGGAGTGCATGCTCTTCAAAAGTTTCAG 114400
 QY 598 AACACATCCACTGACTGAGCATTCAGGGGCAAGAGGAGATGGCAGGCACATTGTTG 657
 Db 114399 AACACATCCACTGACTGAGCATTCAGGGGCAAGAGGAGATGGCAGGCACATTGTTG 114340
 QY 658 ATTGGGTGAGTTTGGGGAGAAATAGACACACAAAGGTCAACATAACTTCTCTAAATTAACA 717
 Db 114339 ATTGGGTGAGTTTGGGGAGAAATAGACACACAAAGGTCAACATAACTTCTCTAAATTAACA 114280
 QY 718 CTTCCTCCATTCACAAATTCCTCTCCCATTTCTCTCCCTGCTCTTTACTSAKARAAA 777
 Db 114279 CTTCCTCCATTCACAAATTCCTCTCCCATTTCTCTCCCTGCTCTTTACTGAAAATAAA 114220
 QY 778 CCCAGTTTTCTGAAACTATATAAAATACCCCAAGTATGTTTACATAATTTACACCTCAA 837
 Db 114219 CCCAGTTTTCTGAAACTATATAAAAGACC--CAGTATTTTACATAATTTACACCTCAA 114161
 QY 838 AGATTAGAAACCAAGAAATAGAGACC--TTTTCACCCCTTCCGGAAGCAAAAGTCATTATCC 896
 Db 114160 AGATTAGAAACCAAGAAATAGAGACCCTTTTCAACCCCTTCCGGAAGCAAAAGTCATTATCC 114101
 QY 897 CTCCAGCCACGTGTCTCAAACTTCATGATGCATCAGAAATCATCTGGGTGCTTTKAAATTCAA 956
 Db 114100 CTCCAGCCACGTGTCTCAAACTTCATGATGCATCAGAAATCATCTGGGTGCTTTGAAAATTCAA 114041
 QY 957 GATGATTCCTACAGGTACCAATAAATCAACTCAGAAATTCCTCCAGTGGGG--CCAGGGAT 1015
 Db 114040 GATGATTCCTACAGGTACCAATAAATCAACTCAGAAATTCCTCCAGTGGGGCCCGGGAT 113981
 QY 1016 CTGTATTTCTGACAAAGCTCCCAAGGTGATTCTCTTCCCAACAGCATTTTGAAGACTTCAG 1075
 Db 113980 CTGTATTTCTGACAAAGCTCCCAAGGTGATTCTCTTCCCAACAGCATTTTGAAGACTTCAG 113921
 QY 1076 CTCAATGACCTAATCAGAGTCTCGCATTTGCTAATATCTGGTCTCATTTTTTBTCTATAT 1135
 Db 113920 CTCAATGACCTAATCAGAGTCTCGCATTTGCTAATTAACCTGCTCATTTTTTTCATATAT 113861
 QY 1136 ATATATAGTATTTGGTAGAGATGGATTTTGGCCATGTTGCCAGGCTAGTATTCAACT 1195
 Db 113860 ATATATAGTATTTGGTAGAGATGGATTTTGGCCATGTTGCCAGGCTAGTATTCAACT 113801
 QY 1196 CCTAAGCTAAGCAATCTTCTCTCTGCTCCCAAAATGTTGGGATTAACAGGTGTAAGC 1255

Db 113800 CCTAAGCTAAGCAATCTTCCTGTCTCTGCTGCCCTCCCAAAATGTTGGATTACAGGTGAAGC 113741

Qy 1256 CACTGCACCGGCTGATAGTGGTTTCATTACTTACTTCTTCTTGGACACTCTCATCCATT 1315

Db 113740 CACTGCACCGGCTGATAGTGGTTTCATTACTTACTTCTTGGACACTCTCATCCATT 113681

Qy 1316 TTGAAGTAAAGTCTCCCAATTATTATGCTGTTTATAGAACACGCTAAGCATGTCATGTGC 1375

Db 113680 TTGAAGTAAAGTCTCCCAATTATTATGCTGTTTATAGAACACGCTAAGCATGTCATGTGC 113621

Qy 1376 TA---ATGGCCAGTGACATCATATAAGAAAGATGTCATTACTGAATGCTTTCATGCTCTTA 1432

Db 113620 TAATTATGGCCAGTGACATCATATAAGAAAGATGTCATTACTGAATGCTTTCATGCTCTTA 113561

Qy 1433 TAATGATGGTAAGTGGCATGTGATGGGGCTATTAG-CCAGACATCACTCCAAAGAA 1491

Db 113560 TAATGATGGTAAGTGGCATGTGATGGGGCTATTAGCCCCAGACATCACTCCAAAGAA 113501

Qy 1492 TTCCAAACAGATATAGACAAGTGCCTTTAGGGCCAGATCCCTCCCTCAGCTGTTTA 1551

Db 113500 TTCCNN 113441

Qy 1552 CCGAGGAATAGATGTCCTGGGACAAAGTTTCCCTAAGTGAAGTGTGAT 1602

Db 113440 NNN 113390

RESULT 5

AL590622/c

LOCUS

DEFINITION

Human DNA sequence from clone RP11-320F15 on chromosome 10.

Contains the gene for ribonuclease P 30kd subunit, the gene for a nuclear protein similar to CARP, ESTs, STSS and GSSs, complete sequence.

AL590622

AL590622.7 GI:14270159

HTG; CARP; ribonuclease P.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 50111)

Tracey, A.

Direct Submission

Submitted (19-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On May 31, 2001 this sequence version replaced gi:14161205.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C.elegans/wormpep

This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr10

RP11-320F15 is from the library RPOI-11.2 constructed by the group of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-320F15 it may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true left end of clone RP11-236B18 is at 50012 in this sequence. The true right end of clone RP11-103A2 is at 100 in this sequence.

FEATURES

Location/Qualifiers

source

1..50111

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosomes="10"

/clone="RP11-320F15"

/clone_lib="RPOI-11.2"

/complement(1..98)

/note="match: GSS: Em:AQ670367"

complement(1..97)

/note="match: GSS: Em:AQ544636"

complement(1..77)

/note="match: STS: Em:G56439"

match: GSS: Em:AQ311708"

120..34469

/gene="BA320F15.1"

join(120..472,3273..3328,3962..4018,4428..4502,7467..7538,14216..14305,23179..23295,23858..23887,24284..24321,24717..24796,28974..29381)

/gene="BA320F15.1"

/products="BA320F15.1.1 (ribonuclease P (30kd) (RPP30))"

/note="match: CDNAS: Em:U77655 Em:AK004137 Em:U95123"

match: ESTs: Em:BE572689 Em:BE378859 Em:BE137893 Em:AA920575 Em:BG106820 Em:BF681617 Em:BF248001 Em:BF031745 Em:AA854455 Em:AI359795"

/evidence="not experimental"

join(391..472,3273..3328,3962..4018,4428..4502,7467..7538,14216..14305,23179..23295,23858..23887,24284..24321,24717..24796,28974..29381)

/gene="BA320F15.1"

/note="match: proteins: Tr:P78346 Tr:O88796"

/codon_start=1

/evidence="not experimental"

/products="BA320F15.1.1 (ribonuclease P (30kd) (RPP30))"

/protein_id="CAC70100.1"

/db_xref="GI:15717970"

/translation="MAVFADLDLRAGSDLKALRGLVTAHLGYSVVAINHIVDFKEK KQIEKPVAVSELFTLPVQKSKPIKILTRLTIVSDPHCNVLRATSSRLRLYDV VAVFPKTEKLPHIACHLDVLCVITVTEKLPFYKRPINVAIDRLGLAFELYSPAI KQSTWRRYTISSALNMQIKGNKVIISAAERPLRPIGPDVANGVLGLFGLSESDAK AAVSTNCRAALHGETRKTAFGIISTVKKPRPSEGEDECLPASKKAKCEG"

689..800

/note="MIR repeat: matches 30..132 of consensus"

1443..1488

/note="L2 repeat: matches 2459..2502 of consensus"

1761..2155

/gene="BA320F15.1"

/note="match: GSS: Em:AQ805524"

1907..2434

/gene="BA320F15.1"

/note="match: GSS: Em:AQ775672"

1926..2208

/gene="BA320F15.1"

/note="match: GSS: Em:B92129"

2349..2482

/note="L2 repeat: matches 2557..2696 of consensus"

2339..3167

/note="MIR repeat: matches 13..262 of consensus"

complement(4176..4649)

/note="match: GSS: Em:AQ626500"

join(<4648..4764,7467..7538,14216..14305,23179..23887,24284..24321,24717..24796,28974..29381)

/gene="BA320F15.1"

/product="BA320F15.1.2 (putative isoform 2)"

/note="match: ESTs: Em:AW939965"

repeat_region

repeat_region

misc_feature

misc_feature

misc_feature

repeat_region

repeat_region

misc_feature

mRNA

REFERENCE AUTHORS TITLE JOURNAL

FEATURES source

BASE COUNT 636 a 567 c 542 g 613 t

ORIGIN

Query Match 17.7% Score 367.8; DB 6; Length 2358;

Best Local Similarity 76.4%; Pred. No. 3.6e-73;

Matches 597; Conservative 0; Mismatches 162; Indels 22; Gaps 11;

QY 1289 TCTATTCTTGACCACTCTGATCCATTTTGAAGTAAATAATGCTCCCAATTATTATGCTGTT 1348

DB 1540 TCTCTGTGTCATCACTTCGGCCGTTTGGGGT--AGATCCTCTGATTAGCCTTCAGAT 1597

QY 1349 TTGAACACCGTAAGCATGTGCTGCTA---ATGCCAGTGCATCATAAAAAGAAAGT 1405

DB 1598 TTGAACACCGTGCAGCTGTGGTGCATTAATTATGGCCAGTGCACCATAGATCAAGT 1657

QY 1406 GCATTACTGAATCTTCAATGCTTATATGATGTTAGTGAAGTGCATGCTGGGCGCTA 1465

DB 1658 GCATTACTGAATCTTCAATGCTTATATGATGTTAGTGAAGTGCATGCTGGGCGCTA 1717

QY 1466 TTTAGC-CCAGACATCACTCCAAAGAAATCCAAACAGATATAGACAGTGCCTTTAGGGC 1524

DB 1718 TTTAGTGCAGACATCACTCCAGAAATTCACACAGATAGACAGTGGCACCAGAC 1777

QY 1525 CCAGATCCCTTCCCTCAGCTGTTTACCAGGAATAGGATGCC---TGGGACAAGTT 1581

DB 1778 CCATCTCTTCCCTCGGGCTGATATCCAGAAATAGGATGTCCTCCAAAGCAACTTC 1837

QY 1582 TCCCTTAAGTGAATGTTGATGCTGCTTATCAGAAATATTAAGTGGGCTGTGATA 1641

DB 1838 CCAGCAACTGGAGTGTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1897

QY 1642 TGAGGCACTCACTTTCTTGATA-GGTAGTCATATGAAGCTGACAAAGAA--AAAA 1698

DB 1898 CACAGTGC--TTGATTTTCTGATAGCTGTTAGTATGATGAGCTGACAAAGAAAGAA 1955

QY 1699 AGGCGAGTGTGGTGCATGTCAACAGACAGCTGTCCCTGAC-TCTTGACAAATAGG 1757

DB 1956 AGAGCAGCATGTGGTGCATATTAACAGCAGCTGTCCCTGGCTTCCGATACGTGG 2015

QY 1758 ATGACTTGATCTGAGCGATGTGATCACCACAAAGAAAGTGGCTCTCAATTTCTT 1817

DB 2016 ATGACTCGATTTCTGAGCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2075

QY 1818 CTTGATTTCATATTCAGAGGGTGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1877

DB 2076 CTTGATTTCATATTCAGAGGGTGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2130

QY 1878 ACTGAGTCTGGAATGAAATTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1937

DB 2131 ACTAAGTCTGGAATGAAATTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2190

QY 1938 TTAATTCGTTCCAGGTTGGAAGATATCTACCGGCGCCAGCTGATATTAAGTCAACG 1997

DB 2191 TGAATTCGTTCCAGGTTGGAAGATATCTACCGGCGCCAGCTGATATTAAGTCAACG 2249

QY 1998 GTGTGAGGGGCCCCAGGCGCAACTCCAGGATTCCTTC-CACACAGAAACATAC 2056

DB 2250 GTGTGAGGGGCCCCAGGCGCGAGTTCCAGGGGTTTCCTCCACAGAGAAACATAG 2309

QY 2057 A 2057

DB 2310 A 2310

AC119234 142902 bp DNA linear HTG 11-JUN-2003

Mus musculus clone RP24-211P24, *** SEQUENCING IN PROGRESS ***, 2 ordered pieces.

AC119234 GI:31581760

HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 142902)

Birren, B., Nusbaum, C. and Lander, E.

Mus musculus, clone RP24-211P24

Unpublished

2 (bases 1 to 142902)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Collings, S., Collins, S., Collamore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Laocque, K., Lamazares, R., Landers, T., Lehotsky, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (25-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 142902)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spenser, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (11-JUN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jun 11, 2003 this sequence version replaced gi:31455706.

All repeats were identified using RepeatMasker:

Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valae, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 229640)
Worley, K.C.

Direct Submission
Submitted (09-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 229640)

Rat Genome Sequencing Consortium.
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 15, 2002 this sequence version replaced gi:23101653.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GNAG
Center clone name: CH230-140118
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 215126 bases at least Q40
Consensus quality: 217436 bases at least Q30
Consensus quality: 218866 bases at least Q20
Estimated insert size: 214300; sum-of-contigs estimation
Quality coverage: 10x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 215467: contig of 215467 bp in length
* 215468 215567: gap of unknown length
* 215568 218782: contig of 3215 bp in length
* 218783 218882: gap of unknown length
* 218883 220359: contig of 1477 bp in length
* 220360 220459: gap of unknown length
* 220460 221614: contig of 1155 bp in length
* 221615 221714: gap of unknown length
* 221715 223060: contig of 1346 bp in length

* 223061 223160: gap of unknown length
* 223161 224771: contig of 1611 bp in length
* 224772 225871: gap of unknown length
* 225872 225886: contig of 1015 bp in length
* 225887 225986: gap of unknown length
* 225987 228180: contig of 2194 bp in length
* 228181 228280: gap of unknown length
* 228281 229640: contig of 1360 bp in length.

FEATURES
source
1..229640
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-140118"
1..1268
/note="wgs_end_extension"
clone_end:T7
4765..7050
/note="wgs_end_extension"
clone_end:T7
/note="wgs_end_extension"
clone_end:T7
complement(7251..8156)
/note="clone_boundary"
clone_end:T7
site:
end_sequence:RWAFX57TJB"
complement(7252..8118)
/note="clone_boundary"
clone_end:T7
site:
end_sequence:RWAFX57TJC"
213542..215467
/note="wgs_end_extension"
clone_end:T7
215568..216791
/note="wgs_end_extension"
clone_end:T7
BASE COUNT 61330 a 47195 c 48097 g 63316 t 9702 others
ORIGIN

Query Match 17.6%; Score 365; DB 2; Length 229640;
Best Local Similarity 77.3%; Pred. No. 1.4e-72;
Matches 601; Conservative 0; Mismatches 150; Indels 26; Gaps 12;

QY 1289 TCTATTTCTTGACCACTCTGATCCATTTTGAAGTAAAAATGCTCCAAATATTATGCTGTT 1348
Db 141439 TGTGTCCTGGACCACTCTGACCCATTTTGGAGTCAATAT--TCCGATTAGCTTCTGTT 141496

QY 1349 TTAGAACACGGTAAGCATGTCATGCTAATGCCAGTGCACATCAATAAAGAAAGTGCA 1408
Db 141497 TTAGAGCAC---ATGCCATGCACATAATATGCCAGTGCACACCATATAAAGTAAAGTGCA 141552

QY 1409 TTACTGAATGCTTTCAATGCTTTATATCATGTAAGTGGCATGTCTATGGGCGCTATTTT 1468
Db 141553 TTACTGAATGCTTTCAATGCTTTATATCATGTAAGTGGCGTGTCTATGGGCGCTATTTT 141612

QY 1469 AGCCCC-AGACATCACTCCAAAGAATTCCAAACAGATATAGACAAGTGCCTTTTAGGGCCCCA 1527
Db 141613 AGCCCTGGACATCACTCCAGAGAAATTTCTGAACAGATATAGACAAGTGTCCACCAAGACCCA 141672

QY 1528 GATCCCTTCCCTCAGGCTGTTTA-CCAGGGAATAGGATGCTCTGGGCAAGTTTCCC- 1585
Db 141673 -----CTTCCCCCGGGCTGTTTATTTCCAGGAATAGGATGTCCCAAGCAACACTTCCA 141727

QY 1586 --CTAAGTGAAGTGTGTGAAGTGTGCTTTATCAGAAAGATATTACTGGGGTGTGTATG 1643
Db 141728 GGCCAACTGGAGTGTGTGATAGCCCGCATTTATCAGAAAGATATTGCCGTAGTGTGATGCA 141787

QY 1644 TAGGGCATCTACATTTTCTTTGATAGTGTCTATATGAAAGCTGACAAAGAA--AAAAAGG 1701
Db 141788 CAATGCTTGCAC-TTTCCTGTAGGTTAGTATACAGAGCTGACAGAGAGGAGAAAGG 141846

QY 1702 GAGTGTATGTGTGCAATCTCAACAGACAGCTGTCCCTGTGACTTTTGACAAATAGGATGA 1761
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

Db 141847 GCAGGATGTGCTGCAATATGAAACAGCGAGCTGTCTCCCTGGCTTCCGATAGTAGGATCA 141906
QY 1762 CTTGCTATTCCTGAGGATGTGATACACCAAGGAATGGCCCTCTCACTTTCTTCCTG 1821
Db 141907 CTTGCTATTCCTGAGGATGTGATACACCAAGGAATGGCCCTCTCACTTTCTTCCTG 141966
QY 1822 ATTACATATTCAGCAGGCTTAGCTTGTCTCCTCCCTCTTCAGCTTCCAGACACACTG 1881
Db 141967 ATTGCGACA--CACACGCGCAGCTTGTCAT---CTCCCTCTGGCTTCCAGACACTA 142021
QY 1882 ACTCTGGAATGAAATTCACCTGCTCTGAGTTGGCTCTTAATGGGGCGGAGTGTAC 1941
Db 142022 AGCTTGGAAATGAAATTCACCTGCTCTGAAATTTGGCCACCGAGGAGCGGGTGTGAC 142081
QY 1942 TTCGGTTCCAGGTTGGAAGATTATCTACCCGGCCCGCAGCTATATAGCTGACCGGTGT 2001
Db 142082 TTGGCTTCCAGGCTGGAAGATTATCTACCCAGTCTAGCTATATAAG-AGGCTGGGT 142140
QY 2002 GGAGGGGCCAGCAGGGCCAACTCCAGGGATTCTTC-CACACAGAGAAAAACATACA 2057
Db 142141 GGAGGGGCTCCACAGGCCAGCTCCAGGGTTTCAGCCACAGAGGAGAAAAACATAGA 142197

RESULT 9
AC097115
LOCUS AC097115 238344 bp DNA linear HTG 14-NOV-2002
DEFINITION Rattus norvegicus clone CH230-26A2, *** SEQUENCING IN PROGRESS ***,
2 unordered pieces.
AC097115
VERSION AC097115.6 GI:24956605
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 238344)
Muzny, D., Marie, M., Metker, M., Lee, A., Branzon, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karkathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshewa, L., Louised, H., Lozada, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Nwakoelamen, O., Okwou, G., Olarnpuneagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,

```

FEATURES
source

```

Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,
Sneed, A., Sodergren, A., Song, X.-Z., Sorelle, R., Soosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vega, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 238344)
Worley, K. C.
Direct Submission
Submitted (11-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 238344)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (14-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 14, 2002 this sequence version replaced gi:22855482.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GQJ
Center clone name: CH230-26A2
----- Summary Statistics
Assembly program: Phrap, version 0.990329
Consensus quality: 222569 bases at least Q40
Consensus quality: 226638 bases at least Q30
Consensus quality: 229427 bases at least Q20
Estimated insert size: 233018; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 2 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
* 1 236546: contig of 236546 bp in length
* 236547 236646: gap of unknown length
* 236647 238344: contig of 1698 bp in length.
Location/Qualifiers
1..238344
/organism="Rattus norvegicus"
/mol_type="genomic DNA"

```



```
/db xref="taxon:10116"
/clone="CH230-26A2"
1..1081
/notes="wgs contig"
BASE COUNT 63454 a 47381 c 49225 g 71429 t 6855 others
ORIGIN
Query Match 17.6%; Score 365; DB 2; Length 238344;
Best Local Similarity 77.3%; Pred. No. 1.4e-72;
Matches 601; Conservative 0; Mismatches 150; Indels 26; Gaps 12;
QY 1289 TCTATTCTTGACACCTCATCTCATTTTCAAGTAAATAAGCTCCCAATTATTATGCTGTT 1348
Db TGTGCTCTGACACACTCTGACCAATTTGAGGTCAATAT--TCCGATTAGCCTTCTGTT 27125
QY 1349 TTAGAACACGGTAAGCATGTCTATGCTGAATGCGCAGTGACATCATATAAAGAAAGAGTGCA 1408
Db 27126 TTAGAGCAC---ATGCCATGCCTAATTTATGCGCAGTGACACATAAAGTAAAGTGCA 27181
QY 1409 TTACTGAATGCTTCAATGCTTATATGATGTGAAGTGGCATGTGATGGGCGCTATT 1468
Db 27182 TTACTGAATGCTTCAATTTCTCATATGATGTGAAGTGGCGTGTGATGGGCGCAATTT 27241
QY 1469 AGCCC-AGACATCACTCCAAAGAAATCCAAACAGATATAGACAAGTGCCCTTTAGGGCCCA 1527
Db 27242 AGCCCTGGACATCACTCCAGAAATTTGAACAGATATAGACAAGTGTCAACAGACCCA 27301
QY 1528 GATCCCTTCCCTCAGGCTGTTTA-CCAGGGAAATAGGATGCTCTGGGACAAAGTTTCCC- 1585
Db 27302 -----CTTCCCGGGCTGTTTATTTCCAGGAATAGGATGCTCCCAAGCACTTCCA 27356
QY 1586 --CTAAGTGAAGTGTGATAAGTCTGTTATATAGAAAGATATTACTGGGGTGTGATG 1643
Db 27357 GGCCTGAGTGTGCTGATAAGCCAGTTATCAGAAAGATATTGCCGTGAGTGTGATGCA 27416
QY 1644 TAGGGCATCTACATTTTCTGATAGGTAGTCATATGAAAGCTGACAAAGNA--AAAAAG 1701
Db 27417 CAATGCTTGAC--TTTCTGATAGGTAGTCATACGAAAGCTGACAGAGAGAGAAAGG 27475
QY 1702 GCAGTGTGTTGGTCAATGTCAACAGACAGCTGCTCCCTGACTCTTGACAAATAGGATGA 1761
Db 27476 GCAGCATGTGTTGCAATGATGACAGGAGCTGCTCCCTGCTTCCGATAGTATGATGA 27535
QY 1762 CTTGCATTTGCTGAGCGATGATGATCAACAAGGAATGGCCCTCTCACATTTCTTCTTG 1821
Db 27536 CTCGCATTTGCTGGCGGTGTGGTCACTGCAATGGAATGGCCCTCTCACATTTCTTCTTG 27595
QY 1822 ATTCATATTTGACGAGGTAGCTTGTCTCTCCCTCCCTCTTTCAGCTTCCGACACTG 1881
Db 27596 ATTCGCACA--CACCACGGCCAGCTTGTCTAT---CTCCCTCTTGGGCTTCCGACACTA 27650
QY 1882 AGTCTGGAATGAAAATTCACTGCTCTGAGTTGGCTCCTAATGGGGGGGAGTGTGTAC 1941
Db 27651 AGTCTGGAATGAAAATTCACTGCTCTGAGTTGGCCACAGCGGAGGAGGGGTGTGAC 27710
QY 1942 TTCGGTTCCAGGTTGGAAGATTATCTCACCGGCGCCAGCTATATAAGCTGACCGGTTG 2001
Db 27711 TTGGCTTCCAGGCTGGAAGATTATCTCACCCAGTCTTACCTATATAAG-AGGCTGGGT 27769
QY 2002 GAGGGGGCCAGCAGGGCCACTCCAGGATTCCTTC-CACGACAGAAAACATACA 2057
Db 27770 GAGGGGGCTCCACAGGGCGAGCTCCAGGGGTTTCAGCCACAAGAGGGAACATAGA 27826
```

RESULT 10

```
AF478692 723 bp DNA linear ROD 10-OCT-2002
LOCUS
DEFINITION Mus musculus cardiac ankyrin repeat protein (Carp) gene, promoter
region and partial cds.
ACCESSION AF478692
VERSION AF478692.1 GI:19110906
KEYWORDS Mus musculus (house mouse)
```

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 723)
AUTHORS Maeda,T., Sepulveda,J., Chen,H.H. and Stewart,A.F.R.
TITLE alpha1-Adrenergic activation of the cardiac ankyrin repeat protein
gene in cardiac myocytes
Gene 297 (1-2), 1-9 (2002)
AUTHORS Maeda,T., Sepulveda,J. and Stewart,A.F.R.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-2002) Cardiovascular Institute, University of
Pittsburgh, 200 Lothrop Street, Pittsburgh, PA 15213, USA

FEATURES

Location/Qualifiers
source 1..723
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="FVB"
/db_xref="taxon:10090"
1..>723
/genes="Carp"
1..720
/genes="Carp"
/note="contains promoter and 5' UTR"
<721..>723
/genes="Carp"
/products="cardiac ankyrin repeat protein"
721..>723
/genes="Carp"
/codon_start=1
/product="cardiac ankyrin repeat protein"
/protein_id="AAL85342.1"
/db_xref="GI:19110907"
/translation="M"
BASE COUNT 187 a 185 c 180 g 171 t
ORIGIN

Query Match 16.4%; Score 340.8; DB 10; Length 723;
Best Local Similarity 74.0%; Pred. No. 5.2e-67;
Matches 527; Conservative 0; Mismatches 172; Indels 13; Gaps 7;

QY 1352 GAACACGGTAAGCATGTCATGTGCTA---ATGCCAGTGACATCATATAAAGAAAGTGCA 1408
Db 1 GAACACGGTGAGCCTCTGGTGACATAATATGSCCAGTGACACCATAGATGCAAGTCGA 60
QY 1409 TTACTGAATGCTTTCAATGCTTATATGATGTAAGTGGCATGTGATGGGCGCTATT 1468
Db 61 TTACTGAATGCTTTCAATTTCTCTAATGCTGATGGCATGTGCACAGGGCCATTTT 120
QY 1469 AGC-CCAGACATCACTCCAAAGAAATTCACACAGATATAGACAAGTGCCTTTAGGGCCCA 1527
Db 121 AGCTGCAGACATCACTCCAGAGAAATTCACACAGATAGAGACAAGTGGCACCCAGACCCA 180
QY 1528 GATCCCTTCCCTCAGGCTGTTTACCAGGGAATAGATGTCTCTGGGACAAGTTTCCCT 1587
Db 181 TCTCTTCCCTCGGCTGATTATCCCCAGAAATAGATGTCCCAAGCAACTTCCAGCC 240
QY 1588 AAGTGAAGTGTGATAAGTCTGCTATCAGAAAGATATTACTGGGGGTGTGATGTAGG 1647
Db 241 AACTGATCGGTGATGATGTCAGATTATCAGAAAGATATGGCTGTAAAGTGTGATCGACAGT 300
QY 1648 GCATCTACATTTTCTTGATAGGTAGTCATATGAAAGCTGACAAAGAAAGAAAGGCGAGTG 1707
Db 301 GCTTGA-TTTTCTTGATAGTTAGTCATATGAGACTGCAAAAGAAAGAAAGAGCAGC 359
QY 1708 ATGTGTGCAATGTCAACAGACAGCTGTCCCTGAC-TCTTGACAAATAGGATGACTTGC 1766
Db 360 GATGTGTGCAATATTAAACAGGAGCTGTCCCTCGCTTCCCGATAGCTGGATGACTCGC 419
QY 1767 ATTGCTGAGCGATGTGATCACCACCAAGGAATGGCCCTCTCACATTTCTTCTGATTCA 1826
Db 420 ATTGCTGAGCGGTGTGCTCACTGCCCCAAGGAATGACCTCTCACATTTCTTCTGATTGC 479

Qy	1827	CATATTCAGCAGGGTTAGCTTGCTCCCTCCCTCTTTCAGCTTCCAGACACTGAGTCT	1886
Db	480	CATACGCGCGG-----CCAGCTGTGTCATCTCCCTCTTGGGCTTCCAGACACTAAGTCT	534
Qy	1887	GGAAATGAAATTCACCTGCCTCTGAGTTGGCTCCTAATGGGGGGGAGTGTACTTCGG	1946
Db	535	GGAAATGAAATTCACCTGCCTCTGAAATGGCCAATGGTGGGGGAGGGGTGACTTGGC	594
Qy	1947	TTCCCGAGTTGGGAAGATTATCTCACCCGGGCCAGCTATATAAGCTGACCGGTGTGGAGG	2006
Db	595	TTCCCGAGCTGGGAAGATTATCTCACCCAGCGCTAGCTATATA--CGGCTGGTGTGGAGG	653
Qy	2007	GGCCCGACGAGGGGCCAACTCCAGGGATTCCTTC-CACGACAGAAAAACATACA	2057
Db	654	GGCTCCACAGGGCCAGTTCCAGGGGTTTCATCCACAGAGAGAAAAACATAGA	705
RESULT 11			
LOCUS	AX322775	Sequence 19 from Patent WO0192567.	1901 bp DNA linear PAT 07-JAN-2002
DEFINITION	AX322775		
ACCESSION	AX322775		
VERSION	AX322775.1	GI:18093755	
KEYWORDS		unidentified	
SOURCE		unidentified	
ORGANISM		unclassified.	
REFERENCE	1		
AUTHORS	Bunk,D., Reuner,B., Beck,J. and Henkel,T.		
TITLE	Novel target genes for diseases of the heart		
JOURNAL	Patent: WO 0192567-A 19 06-DEC-2001;		
	Medigene AG (DE)		
FEATURES	Location/Qualifiers		
source	1..1901		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
BASE COUNT	592 a 378 c 460 g 471 t		
ORIGIN			
Query Match 11.0%; Score 228; DB 6; Length 1901;			
Best Local Similarity 97.6%; Pred. No. 2.3e-41;			
Matches 242; Conservative 0; Mismatches 5; Indels 1; Gaps 1;			
Qy	1828	ATATTACGACAGGGTTAGCTTGT-CCTCCCTCCCTCTTTCAGCTTCCAGACACTGAGTCT	1886
Db	2	AAAAACAGCAGGGTTAGCTTGCTCCCTCCCTCTTTCAGCTTCCAGACACTGATCT	61
Qy	1887	GGAAATGAAATTCACCTGCCTCTGAGTTGGCTCCTAATGGGGGGGAGTGTACTTCGG	1946
Db	62	GGAAATGAAATTCACCTGCCTCTGAGTTGGCTCCTAATGGGGTGGAGTGTACTTCGG	121
Qy	1947	TTCCCGAGTTGGGAAGATTATCTCACCGGCCAGCTATATAAGCTGACCGGTGTGGAGG	2006
Db	122	TTCCCGAGTTGGGAAGATTATCTCACCGGCCAGCTATATAAGCTGACCGGTGTGGAGG	181
Qy	2007	GGCCCGACGAGGGGCCAACTCCAGGGATTCCTTCACGACAGAAAAACATACAAGACTCCTT	2066
Db	182	GGCCCGACGAGGGGCCAACTCCAGGGATTCCTTCACGACAGAAAAACATACAAGACTCCTT	241
Qy	2067	CAGCCAAC 2074	
Db	242	CAGCCAAC 249	
RESULT 12			
LOCUS	BD094076	Sequence 19 from Patent WO0192567.	1901 bp DNA linear PAT 27-AUG-2002
DEFINITION	BD094076	Shear stress-responsive DNAs.	
ACCESSION	BD094076		
VERSION	BD094076.1	GI:22639664	
KEYWORDS	WO 0125427-A/37.		
SOURCE	Hom sapiens (human)		

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1901)
 AUTHORS Myers, R.W.
 JOURNAL Unpublished (1996)
 COMMENT
 Contact: Richard M. Myers
 Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine
 Department of Genetics, M-344, Stanford, CA 94305, USA
 Tel: 4157259687
 Fax: 4157259689
 Email: myers@shgc.stanford.edu
 Primer A: GGCATTTTGAAGGCATGG
 Primer B: CCAGATGGATCATGAAGG
 STS size: 222
 PCR Profile:
 Initial incubation: 94 degrees C for 90 seconds
 Denaturation: 94 degrees C for 15 seconds
 Annealing: 62 degrees C for 23 seconds
 Polymerization: 72 degrees C for 30 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9600
 Protocol:
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 Taq Polymerase: 0.05 units/ul
 Total Vol: 10 ul
 Buffer:
 MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 20 mM
 pH: 8.3
 Prepared with primer pairs provided by Sandoz, derived from X83703
 -- Washington University/Merck EST sequence.
 Location/Qualifiers
 1..1901
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="10"
 1246..1467
 primer_bind 1246..1467
 primer_bind 1246..1467
 complement(1447..1467)
 BASE COUNT 592 a 378 c 460 g 471 t
 ORIGIN
 Query Match 11.0%; Score 228; DB 11; Length 1901;
 Best Local Similarity 97.6%; Pred. No. 2.3e-41;
 Matches 242; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
 QY 1828 ATATTACAGCAGGGTTAGCTTGT-CCTCCCTCCCTCTTTACGCTTCCACAGACACTGAGTCT 1886
 Db 2 AAAAAACAGCAGGGTTAGCTTGTGCTCCCTCCCTCTTTACGCTTCCACAGACACTGATTC 61
 QY 1887 GGAATGAAATTCACCTGCCTCTGAGTTGGCTCCTTAATGGGGCGGGAGTGTCTACTTCGG 1946
 Db 62 GGAATGAAATTCACCTGCCTCTGAGTTGGCTCCTTAATGGGGGTGGGAGTGTCTACTTCGG 121
 QY 1947 TTCCCCAGGTTGGAAGATTATCTCACCCGGCCCCAGCTATATAAGCTGACCCGGTGTGAGG 2006
 Db 122 TTCCCCAGGTTGGAAGATTATCTCACCCGGCCCCAGCTATATAAGCTGACCCGGTGTGAGG 181
 QY 2007 GGCCCGCAGGGCCAACTCCAGGGATTCTCTTCCACGACAGAAAAACATACAAGACTCCTT 2066
 Db 182 GGCCCGCAGGGCCAACTCCAGGGATTCTCTTCCAGCAGAGAAACATACAAGACTCCTT 241

```

QY      2067 CAGCCAAC 2074
Db      242 CAGCCAAC 249

RESULT 15
AF131883
LOCUS   1940 bp mRNA linear MAM 30-NOV-1999
DEFINITION Oryctolagus cuniculus CARP mRNA, complete cds.
ACCESSION AF131883
VERSION   AF131883.1 GI:6478316
KEYWORDS .
SOURCE   Oryctolagus cuniculus (rabbit)
ORGANISM Oryctolagus cuniculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1 (bases 1 to 1940)
AUTHORS   Aihara,Y., Kurabayashi,M., Arai,M., Kedes,L. and Nagai,R.
TITLE     Molecular cloning of rabbit CARP cDNA and its regulated expression
          in adriamycin-cardiomyopathy
JOURNAL   Biochim. Biophys. Acta 1447 (2-3), 318-324 (1999)
MEDLINE   20011295
PUBMED    10542334
REFERENCE 2 (bases 1 to 1940)
AUTHORS   Aihara,Y.
TITLE     Direct Submission
JOURNAL   Submitted (26-FEB-1999) Second Department of Internal Medicine,
          Gunma University School of Medicine, 3-39-15 Showa-machi, Maebashi
          371-8511, Japan
FEATURES             Location/Qualifiers
     source           1..1940
                     /organism="Oryctolagus cuniculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9986"
                     219..1178
     CDS              /codon_start=1
                     /product="CARP"
                     /protein_id="AAF13817.1"
                     /db_xref="GI:6478317"
                     /translation="MMVLKVELVTGKSSNGETGEFLPDDFRDQGYEAAVTSEKQED
                     LKTLPAHVSLAEQWREKQLEAEKKKLEQSKLENLELEIIOLKKRKYRK
                     KYPVAKPEPEIITEPVDVRFKALLENKLAIVVEKFLSDQNNPDVCEYKRTALHRA
                     CLEGLAIVKELMEAGAOIEFDMLESTALHWACRGNLEVLKLLINKAKISARDKL
                     LSTALHVAVRTGHCYCAEHLIACEADLNKADREGDTPLDHDAVRLNRYKMRLLINMGA
                     DUTIKNSAGKTPMDLVLNWQNGTKAIFDLSKENSYKTSKIATP"
BASE COUNT 623 a 391 c 479 g 447 t
ORIGIN

Query Match 7.4%; Score 154.4; DB 4; Length 1940;
Best.Local Similarity 84.9%; Pred. No. 1.3e-24;
Matches 185; Conservative 0; Mismatches 31; Indels 2; Gaps 1;

QY      1859 CCTCTTCAGCTTCCCAGACACTGAGTCTGGAAATGAAAAATTCACCTGCCTCTGAGTTGGCT 1918
Db      1 CCTCTTGGGCTTCCCAGACACTGAGTCTGGAAATGAAAAATTCACCTGCCTCTGAGTTGGCT 60

QY      1919 CTT--AATGGGGCGGGAGTGTACTTCGGTTCCAGGTTGGAAGATTATCTACCCGGC 1976
Db      61 ACTGGCAGGGGGCAGGGCTGTACTTGGGTTCCAGGTTGGAAGATTATCTACCCAGC 120

QY      1977 CCCAGTATATAGCTGACGGGTGTGGAGGGCCCGAGCGGCCCACTCCAGGGATTCT 2036
Db      121 CCAACTATATAGCGGAGCTGTGGAGGGGACCGAGCGGGCCATCTCCAGGGATTCT 180

QY      2037 TCCACGACAGAAAAACATACAAAGACTCTCTTCAGCCCAAC 2074
Db      181 TCCACAGGAGAAACACACAGCGCACTCCAGCCCAAC 218

```

Search completed: November 14, 2003, 11:17:48
 Job time : 7542.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 03:46:02 ; Search time 535.347 Seconds
(without alignments)
10457.963 Million cell updates/sec

Title: US-10-005-337A-2
Perfect score: 2074
Sequence: 1 ctgcacgaagttacttaag.....acaagactcttcagccaac 2074

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2071.6	99.9	2074	ABV73021	Human CARP protein
2	367.8	17.7	2358	ABV73020	Mouse CARP protein
3	322.6	15.6	2247	AAI10406	Murine cardiac ank
4	229.6	11.1	1988	AAI93508	Human polynucleoti
5	228	11.0	1901	AAH02910	Human shear stress
6	228	11.0	1901	AAH02910	Human CAAS8676 pro
7	144.4	7.0	1889	AAH94903	Human DNA sequence
8	144.4	7.0	1889	ABX77648	Differentially exp

9	97.2	4.7	34488	22	AAF97854	Human neuroblastom
10	97.2	4.7	121162	21	AAK66548	Human kinesin-like
11	95.6	4.6	7141	22	AAK68280	Human immune/haema
C 12	95.4	4.6	6040	22	AAS28595	Genomic sequence #
C 13	94	4.5	4181	22	AAK84643	Human immune/haema
C 14	94	4.5	7141	22	AAK68282	Human immune/haema
C 15	93.6	4.5	52845	22	AAK71437	Human immune/haema
C 16	93.4	4.5	29329	22	ABA18026	Human nervous syst
C 17	93.4	4.5	29329	22	ABA20511	Human nervous syst
C 18	93.4	4.5	29329	22	AAK70791	Human immune/haema
C 19	93.4	4.5	29329	22	AAK78512	Human immune/haema
C 20	93.2	4.5	32220	22	AAK78512	Human immune/haema
21	93.2	4.5	32220	22	AAK78512	Human immune/haema
22	92.8	4.5	90220	24	ABX60556	CDNA encoding nove
23	92.8	4.5	90220	24	ABX60556	CDNA encoding nove
24	92.6	4.5	19408	22	AAZ50905	Human TBC-1 partia
25	92.6	4.5	19408	22	AAZ50905	Genomic sequence #
26	92.6	4.5	19408	22	AAK87230	Human immune/haema
27	92.4	4.5	4678	22	AAK90644	Human digestive sy
28	92.2	4.4	66804	24	ABK87050	Human nervous syst
29	92	4.4	355	22	AAK60063	Human transporter
30	91.4	4.4	3953	22	AAK85377	Human immune/haema
31	91.4	4.4	3985	22	AAK85378	Human immune/haema
C 32	91.2	4.4	3073	24	ABQ77794	Human CGMP-inhibic
C 33	91	4.4	26390	22	AAK65971	Human immune/haema
C 34	91	4.4	56737	24	ABX69895	Human immune/haema
C 35	90.6	4.4	5248	22	AAK65595	Human hypoxanthine
C 36	90.6	4.4	5249	22	AAK65596	Human immune/haema
C 37	90.6	4.4	5249	22	AAK65597	Human immune/haema
C 38	90.6	4.4	118384	25	ABX56555	Human immune/haema
C 39	90.4	4.4	2778	21	AAK81712	Human autoimmu d
C 40	90.4	4.4	19820	22	AAK36348	Human secreted pro
C 41	90.4	4.4	19820	25	ABX59336	Human musculoskele
C 42	90.2	4.3	36785	22	AAK82208	CDNA encoding nove
43	90	4.3	345	22	AAK76845	Human immune/haema
44	90	4.3	22230	24	ABK84349	Human immune/haema
45	89.8	4.3	86592	25	ABZ22285	Human CDNA differe
						Human D-amino acid

ALIGNMENTS

RESULT 1
ABV73021
ID ABV73021 standard; DNA; 2074 BP.
XX
AC ABV73021;
XX
DT 08-JAN-2003 (first entry)
XX
DE Human CARP protein coding sequence upstream DNA fragment.
XX
DE Cardiac ankyrin repeat protein; CARP; cardiac; immunosuppressive;
XX
KW antiinflammatory; gene therapy; antisense gene therapy; human; ds.
XX
OS Homo sapiens.
PN WO200246220-A2.
XX
XX 13-JUN-2002.
XX
PF 05-DEC-2001; 2001WO-EP15412.
XX
XX 07-DEC-2000; 2000US-251582P.
XX
XX (AVET) AVENTIS PHARMA SA.
XX (REGC) UNIV CALIFORNIA.
XX (BENO) BENOIT P.
XX
PI Schwartz B, Branellec D, Chien K;
XX WPI; 2002-740642/90.
XX

PT New promoter sequence derived from a portion upstream of the coding
PT sequence of a gene for Cardiac Ankyrin Repeat Protein, for controlling
PT the level and specificity of expression of a transgene in cardiac
PT muscle cells -

PS Claim 5; Fig 2; 48pp; English.

XX The invention relates to a polynucleotide (I) comprising a fragment of a
CC sequence upstream of the coding part of the gene for the Cardiac Ankyrin
CC Repeat Protein (CARP). (I) is capable of inducing a specific expression
CC in vivo of a gene operably linked to (I), in cardiac cells. (I) or a
CC vector (Iib) comprising (I) is useful for the manufacture of a medicament
CC intended for the treatment of cardiac insufficiency, cardiac hypertrophy
CC and hypoxia, and for preventing rejection during cardiac transplant. An
CC expression cassette under the control of (I) is useful for encoding a
CC protein or RNA which is capable of activating the growth of cardiac
CC cells, reducing or suppressing an immune response, inducing angiogenesis,
CC correcting muscle contractility, cardiac hypertrophy, cardiac
CC insufficiency and myocarditis. (Iib) is useful for expressing a gene of
CC therapeutic interest in vivo, by isolating (Iib) and introducing (Iib) in
CC the cardiac tissue, under conditions so that the gene of interest is
CC expressed. (I), the vectors and the compositions are useful in clinical,
CC experimental, therapeutic and diagnostic fields, and in the treatment and
CC prevention of cardiac pathologies. (I) is also useful for generating
CC transgenic animals which constitute models for studying certain cardiac
CC pathologies. The transgenic animals are also useful for screening
CC molecules for their activity on the regulatory sequences of the gene
CC encoding the CARP protein. The present sequence represents the DNA
CC fragment upstream of the coding sequence of a human CARP protein.

XX Sequence 2074 BP; 612 A; 469 C; 416 G; 572 T; 5 other;

Query Match 99.9%; Score 2071.6; DB 24; Length 2074;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2074; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CTGACGCAAGTTACTTAATGTTTTTGGCTCAGCATCTCTCTGTAATAATGAGCATTA	60
DB	1	CTGACGCAAGTTACTTAATGTTTTTGGCTCAGCATCTCTCTGTAATAATGAGCATTA	60
QY	61	GTCTTGCTCCAACTTCGAGGCGATGGACAGCTCTGGGATTTTCATATCCAAAGCCCTTAA	120
DB	61	GTCTTGCTCCAACTTCGAGGCGATGGACAGCTCTGGGATTTTCATATCCAAAGCCCTTAA	120
QY	121	CATCCACAGTCTTCCGCCAAACACTTCTCTCTTAATACCTCCCTCAGTTGGGTGAG	180
DB	121	CATCCACAGTCTTCCGCCAAACACTTCTCTCTTAATACCTCCCTCAGTTGGGTGAG	180
QY	181	GCTCGAACAAGGATACGAAATGGTAGAAAAGTGTCCATGACTACTTCTGACTTA	240
DB	181	GCTCGAACAAGGATACGAAATGGTAGAAAAGTGTCCATGACTACTTCTGACTTA	240
QY	241	GATGAGAGACCAATGAAATAGTATGACTCTGTTTGTCTCAGCAGGACATATACATA	300
DB	241	GATGAGAGACCAATGAAATAGTATGACTCTGTTTGTCTCAGCAGGACATATACATA	300
QY	301	ATAGGAGCTATACAAAGAGATTAGCATGGACTCTGTGCAAGATGACACAAATTTGT	360
DB	301	ATAGGAGCTATACAAAGAGATTAGCATGGACTCTGTGCAAGATGACACAAATTTGT	360
QY	361	GAAACATTCATATATTAATAATAATAATAATAATAATAATAATAATAATAATAATA	420
DB	361	GAAACATTCATATATTAATAATAATAATAATAATAATAATAATAATAATAATAATA	420
QY	421	AAATAGTATAGTCTGTGCTTCTCAAGAAAGCCAGGAGATTTCTTTATTACCC	480
DB	421	AAATAGTATAGTCTGTGCTTCTCAAGAAAGCCAGGAGATTTCTTTATTACCC	480
QY	481	CTTTTAAGATAGATATTAAGGACCGGAAACATATGATACAGAGGTACTGGAGGGTCC	540
DB	481	CTTTTAAGATAGATATTAAGGACCGGAAACATATGATACAGAGGTACTGGAGGGTCC	540
QY	541	CTCTTTGTCAAATGTTTTGTCTTGGGGTGGGAGTCGATGCTCTCTCAAAGTTTCAGAAC	600

Db 1621 GATATTACTGGGGTGTGATATCTAGGGCATCTACATTTTCTTGATAGGTAGTCATATGA 1680
Qy 1681 AAGCTGACAAAGAAAAGGGCAGTGTATGTGGTCAATGTCAACAGACAGCTGTCCCT 1740
Db 1681 AAGCTGACAAAGAAAAGGGCAGTGTATGTGGTCAATGTCAACAGACAGCTGTCCCT 1740
Qy 1741 GACTCTTGACAAATAGGATGACTTGCATTTCTCTGATGATGCTGAGCGATGTGATCAACCAAGGAATG 1800
Db 1741 GACTCTTGACAAATAGGATGACTTGCATTTCTCTGATGATGCTGAGCGATGTGATCAACCAAGGAATG 1800
Qy 1801 GCCTCTCACATTTCTCTGATTTACATATTACAGCAGGGTTAGCTGTCTCTCCCTCCC 1860
Db 1801 GCCTCTCACATTTCTCTGATTTACATATTACAGCAGGGTTAGCTGTCTCTCCCTCCC 1860
Qy 1861 TCTTACGCTCCAGACACTGAGTCTGGAATGAAATTCACCTGCTGAGTTGGCTCC 1920
Db 1861 TCTTACGCTCCAGACACTGAGTCTGGAATGAAATTCACCTGCTGAGTTGGCTCC 1920
Qy 1921 TAATGGGGCGGAGTGTACTTCTCGTTCCAGGTTGGAAGATATCTACCCGGGCCCA 1980
Db 1921 TAATGGGGCGGAGTGTACTTCTCGTTCCAGGTTGGAAGATATCTACCCGGGCCCA 1980
Qy 1981 GCTATATAAGCTGACCGGTGTGAGGGGCCAGCAGGGCCAACTCCAGGGATTCTTTCCA 2040
Db 1981 GCTATATAAGCTGACCGGTGTGAGGGGCCAGCAGGGCCAACTCCAGGGATTCTTTCCA 2040
Qy 2041 CGACAGAAAACATACAAGACTCCTTCAGCCAAC 2074
Db 2041 CGACAGAAAACATACAAGACTCCTTCAGCCAAC 2074

RESULT 2

ABV73020
ID ABV73020 standard; DNA; 2358 BP.

AC ABV73020;

XX 08-JAN-2003 (first entry)

DE Mouse CARP protein coding sequence upstream DNA fragment.

KW Cardiac ankyrin repeat protein; CARP; cardiant; immunosuppressive;
XX antiinflammatory; gene therapy; antisense gene therapy; mouse; ds.

OS Mus musculus.

XX WO200246220-A2.

XX 13-JUN-2002.

XX 05-DEC-2001; 2001WO-EP15412.

XX 07-DEC-2000; 2000US-251582P.

XX (AVET) AVENTIS PHARMA SA.

PA (REGC) UNIV CALIFORNIA.

PA (BENO/) BENOIT P.

XX Schwartz B, Branellec D, Chien K;

XX WPI; 2002-740642/80.

XX New promoter sequence derived from a portion upstream of the coding
PT sequence of a gene for Cardiac Ankyrin Repeat Protein, for controlling
PT the level and specificity of expression of a transgene in cardiac
PT muscle cells

XX Claim 1; Fig 1; 48pp; English.

XX The invention relates to a polynucleotide (I) comprising a fragment of a
CC sequence upstream of the coding part of the gene for the Cardiac Ankyrin
CC Repeat Protein (CARP). (I) is capable of inducing a specific expression
CC in vivo of a gene operably linked to (I), in cardiac cells. (I) or a

CC vector (Iib) comprising (I) is useful for the manufacture of a medicament
CC intended for the treatment of cardiac insufficiency, cardiac hypertrophy
CC and hypoxia, and for preventing rejection during cardiac transplant. An
CC expression cassette under the control of (I) is useful for encoding a
CC protein or RNA which is capable of activating the growth of cardiac
CC cells, reducing or suppressing an immune response, inducing angiogenesis,
CC correcting muscle contractility, cardiac hypertrophy, cardiac
CC insufficiency and myocarditis. (Iib) is useful for expressing a gene of
CC therapeutic interest in vivo, by isolating (Iib) and introducing (Iib) in
CC the cardiac tissue, under conditions so that the gene of interest is
CC expressed. (I), the vectors, and the compositions are useful in clinical,
CC experimental, therapeutic and diagnostic fields, and in the treatment and
CC prevention of cardiac pathologies. (I) is also useful for generating
CC transgenic animals which constitute models for studying certain cardiac
CC pathologies. The transgenic animals are also useful for screening
CC molecules for their activity on the regulatory sequences of the gene
CC encoding the CARP protein. The present sequence represents the DNA
CC fragment upstream of the coding sequence of a mouse CARP protein.

XX Sequence 2358 BP; 636 A; 567 C; 542 G; 613 T; 0 other;

Query Match 17.7%; Score 367.8; DB 24; Length 2358;
Best Local Similarity 76.4%; Pfd. No. 7.3e-76;
Matches 597; Conservative 0; Mismatches 162; Indels 22; Gaps 11;

Qy 1289 TCTATTCTTGACCACCTCTGATCCATTTTGAAGTAAAAATGCTCCAATTATTATGCTGTT 1348

Db 1540 TCTCTCTGTGCATCACTTCGGCCCGTTTGGGGT--AGATCCCTCTGATTAGCCTTCAGAT 1597

Qy 1349 TTAGAACACGGTAAGCATGTCTATGCTA---ATGGCCAGTGACATCATAAAGAAAGT 1405

Db 1598 TTAGAACACGGTAAGCATGTCTATGCTA---ATGGCCAGTGACATCATAAAGT 1657

Qy 1406 GCATTACTGAATGCTTTCAATGCTTTAATGATGTAAGTGGCATGTCATCGGGCCTA 1465

Db 1658 GCACTACTGAATGCTTTCAATGCTTTAATGATGTAAGTGGCATGTCATCGGGCCTA 1717

Qy 1466 TTAGG-CCAGACATCACTCCAAAGAAATCCAAACAGATATAGACAAGTGCCTTTAGGGC 1524

Db 1718 TTAGGCTGCAGACATCACTCCAGAGAAATCCAAACAGATAGACAAAGTGGCACCCAGAC 1777

Qy 1525 CCAGATCCCTTCCCTCAGGCTGTTTACCAGGGAATAGATGTCC---TGGGACAAAGT 1581

Db 1778 CCATCTCTCTCCCTCGGGCTGATTATCCCCAGAAATAGGATGTCCCAAAGCAACACTTC 1837

Qy 1582 TCCCTTAAGTGAAGTCTTCAATGCTGCTTATCAGAAAGATATTACTGGGGTGTGATA 1641

Db 1838 CCAGCCAACTGGAGTCTGATAGTCCAGTTATCAGAAAGATATGCTGTAAGTGTGATG 1897

Qy 1642 TGTAGGGCATCTACATTTTCTTGATA-GGTAGTCTATGAAAGCTGACAAAGAA--AAAA 1698

Db 1898 CACAGTGC--TTGCATTTTCTTGATACGTTAGTATATGAGAGTGCACAAAGAGGAAA 1955

Qy 1699 AGGGCAGTATGTGGTGCATGTCAACAGACAGCTGTCCCTGAC-TCTTGAACAATAGG 1757

Db 1956 AGAGCAGCGATGTGGTGCATATTAACAGGCAAGTGTCCCTGGCTTCCGATACGTTGG 2015

Qy 1758 ATGACTTGCATGTGAGCGATGTGATCACCACCAAGAAATGGCCCTCTCACATTTCTT 1817

Db 2016 ATGACTTGCATGTGAGCGGTGTGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2075

Qy 1818 CCTGATTTACATATTTCAGCAGGGTTAGCTTTGCTCTCCCTCTCTCTCTCTCTCTCTCT 1877

Db 2076 CTTGATTCGCATACGCCGGG-----CCAGCTTGTCTCTCTCTCTCTCTCTCTCTCT 2130

Qy 1878 ACTGAGTCTGGAAATGAAATTCACCTGCTCTGAGTTGGCTCTCTAAATGGGGGGGAGTG 1937

Db 2131 ACTAAGTCTGGAATGAAATTCACCTGCTCTGATTTGGCCACTGGTGGGGGCGAGGGTG 2190

Qy 1938 TTACTTCGGTCCCGAGTTGGAGATTATCTACCCGGGCCCGAGCTATATAGCTGACCG 1997

Db 2191 TGACTTGGTCTCCAGGCTGGAGATTATCTACCCAGGCCCTAGCTATATAA-CGGGCTG 2249

QY 1998 GTGTGGAGGGCCAGCAGGGCCAACTCCAGGGAATTCCTTC-CAGCAGAGAAAAACATAC 2056
AAA10406
DB 2250 GTGTGGAGGGGTCCACAGGGCCAGTTCAGGGGTTCACACAGAGAGAAAAACATAG 2309
QY 2057 A 2057
DB 2310 A 2310

RESULT 3

AAA10406
ID AAA10406 standard; DNA; 2247 BP.

XX
AC AAA10406;

XX 18-JUL-2000 (first entry)

XX Murine cardiac ankyrin repeat protein (CARP) promoter.

XX Cardiac ankyrin repeat protein; CARP promoter; murine; adenovirus vector;
KW cardiac specific; heart disease; gene therapy; ds.

XX Mus musculus.

XX WO200015821-A1.

XX 23-MAR-2000.

XX 10-SEP-1999; 99WO-US20730.

XX 11-SEP-1998; 98US-0099960.

XX (REGC) UNIV CALIFORNIA.

XX Chien KS, Wang Y, Evans S;

XX WPI; 2000-271457/23.

XX Human type-5 recombinant adenovirus vector used for targeted gene
PT therapy for heart disease and evaluating gene function contains a
PT tissue-restricted promoter and inverted terminal repeat sequences -

XX Claim 8; Page 29-30; 33pp; English.

XX The invention relates to a human type-5 recombinant adenovirus vector
CC for achieving cardiac-restricted transcription of a gene of interest.
CC The vector comprises inverted terminal repeat (ITR) sequences from human
CC adeno-associated virus (AAV) type 2 (AAA10404-A10405) and a cardiac
CC tissue- specific promoter. In particular, the promoter is that of the
CC cardiomyocyte- restricted cardiac ankyrin repeat protein (CARP) gene.
CC The adenovirus vector is used for targeted gene therapy for heart
CC disease and for evaluating gene function. Cardiac restricted
CC transcription of a transgene in both neonatal and mature cardiac tissues
CC can be achieved to treat inherited and acquired heart diseases. The
CC vector is suitable for tissue-specific use in vivo and in vitro and
CC provides cardiac restricted transcription. The present sequence
CC represents the murine cardiac ankyrin repeat protein (CARP) promoter.

XX SQ Sequence 2247 BP; 609 A; 549 C; 485'G; 589 T; 15 other;

Query Match 15.6%; Score 322.6; DB 21; Length 2247;
Best Local Similarity 76.3%; Pred.No. 2.5e-65;

Matches 567; Conservative 0; Mismatches 154; Indels 22; Gaps 13;

QY 1325 AAATGCTCCAATTATATGCTGTTTTAGAACACGGTAAAGCATGTCTAT--GTGCTAATGGC 1382
DB 1499 AGATCTCTGATTAGCTTCAGATTTAGAACACGGTGAAGCTGTGCTCACTAATTATGCG 1558

QY 1383 CAGTGCATCATAAAGAGAAAGTGCATTTACTGAATGCTTTCAATGCTTTATATATGATGTT 1442
DB 1559 CAGTGACACCATAGAGTCAAGTGCATTTACTGAATGCTTTCAATTTCTCTAATGCTGTT 1618

QY 1443 AAGTGGCATGTCATGGGGCTATTTTAGCCGACATCACTCCAAAGAAATTCACACAGA 1502

DB 1619 ACGATGGCATGTACAGGGCCATTTTAGCTGCAGACATCATCCAGAGAAATTCACAAACAGA 1678
QY 1503 TATACAGAAGTGCCTTTTAGGGCCAGATCCCTTCCCTCAGGCTGTATTACCCAGGGAATA 1562
DB 1679 TA-GGACAAGTGGCACCAGACCCCATCT-CCTTCCCTCGGGCTGATTATCCCCAAAAATA 1736
QY 1563 GGATGTCC--TGGGACAAAGTTTCCCTAAAGTGAAGTGTGATAAGTCTGCTTATCAGAA 1619
DB 1737 GGATGTCCAAAGCAACACTTCCAGCCAACTGGAGTGTGATAAGTCCAGTTATCAGAA 1796
QY 1620 AGATATTACTGGGGGTGTGATATGATGAGGCATCTACATTTTCTTCTGATA-GGTAGTCATAT 1678
DB 1797 AGATATGCTGTAAAGTGTGATGCACAGTGC--TTGCAATTTTCTTGATACGTTAGTCATAT 1854
QY 1679 GAAAGCTGACAAAGAA--AAAAAGGGCAGTGTGTTGTCATGTCAACAGACAGCTGTC 1736
DB 1855 GAGAGCTGACAAAGAGGAAAAAGAGCAGCGATGT-GTGCAATATTAACAGACAGCTGTC 1913
QY 1737 CCTTGAC--TCTTGACAAATAGGATGATGATTCGATTCCTGAGCGATGTGATCACCACCAAG 1795
DB 1914 CCTTGGCTTCCGATACGTGGGATGACTCGCATTCCTGAGCGGTGTGTCACCTGCAAG 1973
QY 1796 GAAAGCTGCTCCTCACATTTCTTCTGATTCACATATTGACAGGGTGTAGCTTGTCTCTCC 1855
DB 1974 GAATGACCTCTCACATTTCTTCTGATTCGATACGCGCGG-----CCAGCTTGTCTAT 2028
QY 1856 CTCCTCTTTCAGCTTCCAGACACTGAGTCTGGAATGAAAAATTCACCTGCCTCTGAGTTG 1915
DB 2029 CTCCTCTTGGGCTTCCAGACACTAAGTCTGGAATGAAAAATTCACCTGCCTCTGAAATTG 2088
QY 1916 GCTCCTAATGGGGGGGAGTGTACTTCGGTTCCAGGTTTCCAGAGATTATCTCACCCGG 1975
DB 2089 GCCACTGGTGGAGCAGGGGTGTGACTTGGCTTCCAGGCTGGAAGATTATCTCACCCAG 2148
QY 1976 CCCAGCTATATAAGCTGACCGGTGTGGAGGGGCCCCAGAGGGCCAACTCCAGGGATTCC 2035
DB 2149 CCTTA-CTATATAA-CGGGCTGTGTGGAGGGGCTCCACAGGGCCAGTTCCAGGGGTTCA 2206
QY 2036 TTC-CACGACAGAAAAACATACA 2057
DB 2207 TCCACAAGAGAGAAAAACATAGA 2229

RESULT 4

AA193508
ID AA193508 standard; cDNA; 1988 BP.

XX AC AA193508;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 13568.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX -28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;


```
FT misc_feature /product= "Human CAA58676 protein"
FT 198..431 /*tag= b
FT /note= "66268 cDNA fragment"
FT 1604..1753 /*tag= b
FT /note= "S1MC01-1 cDNA fragment"
FT 1302..1306 /*tag= b
FT /note= "AU-rich mRNA decay element"
FT 1391..1401 /*tag= b
FT /note= "AU-rich mRNA decay element"
FT 1415..1423 /*tag= b
FT /note= "AU-rich mRNA decay element"
XX WO200192567-A2.
XX 06-DEC-2001.
XX 30-MAY-2001; 2001WO-EP06165.
XX 30-MAY-2000; 2000US-207400P.
XX (MEDI-) MEDIGENE AG.
XX Bunk D, Reuner B, Beck J, Henkel T;
XX WPI; 2002-122073/16.
XX P-PSDB; AA16633.
XX Identifying a subject at risk for a heart disease e.g. congestive heart
XX failure, dilative cardiomyopathy, heart muscle disease, by quantifying
XX the polypeptide expressed by genes abnormally expressed in heart tissue
XX
XX Claim 2a; Fig 10b; 154pp; English.
XX
XX The patent discloses novel target genes abnormally expressed in heart
XX tissues and their corresponding proteins. The invention also relates to
XX methods for assessing the expression level of these genes. The method
XX is used for testing the predisposition of mammals and preferably humans
XX for a heart disease or for an acute state of such a disease. It is also
XX useful to treat diseases of the heart such as congestive heart failure,
XX dilative cardiomyopathy, hypertrophic cardiomyopathy, ischaemic cardio-
XX myopathy, specific heart muscle disease, rhythm and conduction disorders,
XX syncope and sudden death, coronary heart disease, systemic arterial
XX hypertension, pulmonary hypertension, pulmonary heart disease, valvular
XX heart disease, congenital heart disease, pericardial disease and
XX endocarditis. Sequences of the invention are also used in gene therapy.
XX A transgenic non-human mammal comprising the sequences of the invention
XX are useful for the development for medicaments for the treatments of
XX heart diseases. The present DNA sequence is expressed sequence tag
XX (EST) clone X83703 which encodes CAA58676 protein.
XX
XX Sequence 1901 BP; 592 A; 378 C; 460 G; 471 T; 0 other;
XX
XX Query Match 11.0%; Score 228; DB 24; Length 1901;
XX Best Local Similarity 97.6%; Pred. No. 2.8e-43;
XX Matches 242; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
XX
XX 1828 ATATTGAGGAGTTAGCTTGT-CCTCCCTCCCTTTCAGCTTCCAGACACTGAGTCT 1886
XX 2 AAAAACAGCAGGAGTTAGCTTGTCCCTCCCTCCCTTTCAGCTTCCAGACACTGAGTCT 61
XX
XX 1887 GGAATCAAAATTCACCTGCTCTGAGTTGGCTCTTAATGGGGCGGAGTGTACTTCGG 1946
XX 62 GGAATCAAAATTCACCTGCTCTGAGTTGGCTCTTAATGGGGCGGAGTGTACTTCGG 121
XX
XX 1947 TTCCAGAGTTGGAAGATTATCTCACCCGGCCCCAGCTATATAAGCTGACCGGTGTGGAGG 2006
XX 122 TTCCAGAGTTGGAAGATTATCTCACCCGGCCCCAGCTATATAAGCTGACCGGTGTGGAGG 181
XX
XX Query Match 7.0%; Score 144.4; DB 24; Length 1889;
XX Best Local Similarity 99.3%; Pred. No. 8.9e-24;
XX Matches 145; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1929 GCGGAGTGTACTTTCGGTTCCAGGTTGGAAGATTATCTCACCCGGCCCCAGCTATATA 1988
XX 5 GAGGAGTGTACTTTCGGTTCCAGGTTGGAAGATTATCTCACCCGGCCCCAGCTATATA 64
XX
XX 1989 AGCTGACCGGTGTGGAGGGGGCCAGCAGGGCCAACTCCAGGGATTCCTTCCACGACAGAA 2048
XX 65 AGCTGACCGGTGTGGAGGGGGCCAGCAGGGCCAACTCCAGGGATTCCTTCCACGACAGAA 124
XX
```


QY 2049 AACATACAGAGCTCCTTCAGCCAAC 2074
 Db 125 AACATACAGAGCTCCTTCAGCCAAC 150

RESULT 8

ABX77648
 ID ABX77648 standard; cDNA; 1889 BP.
 XX AC ABX77648;

XX DT 09-APR-2003 (first entry)

XX DE Differentially expressed breast cancer associated cDNA #143.

XX KW Breast cancer; differential gene expression; BC-cDNA;
 KW breast cancer diagnosis; breast cancer monitoring;
 KW breast cancer treatment; breast cancer staging; gene; ss.

XX OS Homo sapiens.

XX PN US2002156263-A1.

XX PD 24-OCT-2002.

XX PF 04-OCT-2001; 2001US-0974298.

XX PR 05-OCT-2000; 2000US-238331P.

XX PA (CHEN/) CHEN H.

XX PI Chen H;

XX DR WPI; 2003-182653/18.

XX PT New cDNAs, which are differentially expressed in (metastatic) breast
 PT cancer useful for diagnosing or staging, breast cancer, or for
 PT monitoring the treatment of breast cancer in an individual -
 XX Claim 1; SEQ ID NO 182; 30pp; English.

XX CC The invention describes a combination of cDNAs (designated BC-cDNAs),
 CC which are differentially expressed in breast cancer. The combination
 CC includes 152 cDNA sequences, or their complements. The protein encoded
 CC by any of these BC-cDNAs is useful for screening several molecules or
 CC compounds to identify at least one ligand that specifically binds the
 CC protein, producing or preparing polyclonal or monoclonal antibodies, or
 CC purifying antibodies from a sample. The antibodies, which specifically
 CC bind the protein differentially expressed in breast cancer is useful for
 CC detecting the expression of a protein in a sample. The BC-cDNAs are
 CC also useful for diagnosing, monitoring the treatment of, or staging,
 CC breast cancer. This sequence represents a differentially expressed
 CC breast cancer associated cDNA.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from the US patent office at
 CC seqdata.uspto.gov/sequence.html?DocID=20020156263.

XX SQ Sequence 1889 BP; 613 A; 354 G; 452 G; 469 T; 1 other;

Query Match 7.0%; Score 144.4; DB 25; Length 1889;
 Best Local Similarity 99.3%; Pred. No. 8.9e-24;
 Matches 145; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1929 GCGGAGTGTTACTTCGGTCCAGGTGGAGATTATCTCACCCGGCCCGCTATATA 1988
 Db 5 GAGGAGTGTTACTTCGGTCCAGGTGGAGATTATCTCACCCGGCCCGCTATATA 64

QY 1989 AGCTGACCGGTGTGGAGGGGGCCAGCAGGCCCAACTCCAGGGATTCTTCCACGACAGAA 2048

Db 65 AGCTGACCGGTGTGGAGGGGGCCAGCAGGCCCAACTCCAGGGATTCTTCCACGACAGAA 124

QY 2049 AACATACAGAGCTCCTTCAGCCAAC 2074
 Db 125 AACATACAGAGCTCCTTCAGCCAAC 150

RESULT 9

AAF97854
 ID AAF97854 standard; DNA; 34488 BP.

XX AC AAF97854;

XX DT 31-MAY-2001 (first entry)

XX DE Human neuroblastoma cell line NB-1 lp36 nucleotide sequence SEQ ID NO:68.

XX KW Human; chromosome 1; lp36; neuroblastoma cell line; NB-1; anticancer;
 KW tumour suppressor; human lp36 homozygosity deletion domain; tumour;
 KW diagnosis; ds.

XX OS Homo sapiens.

XX PN WO200116311-A1.

XX PD 08-MAR-2001.

XX PF 31-AUG-2000; 2000WO-JP05930.

XX PR 31-AUG-1999; 99JP-0245962.

XX PR 09-MAY-2000; 2000JP-0136266.

XX PA (HISM) HISAMITSU PHARM CO LTD.

XX PA (CHIB-) CHIBA PREFECTURE.

XX PI Nakagawara A;

XX DR WPI; 2001-226686/23.

XX PT Human lp36 homozygosity deletion domain from the 36-position of first
 PT chromosome short arm in human neuroblastoma cell lines, applicable e.g.
 PT in gene diagnosis of tumors as well as in developing anti-cancer drugs
 XX Example 8; Page 104-118; 226pp; Japanese.

XX CC The present invention describes a homozygosity deletion domain
 CC co-existing in the 36-position of the first chromosome short arm (lp36)
 CC in human neuroblastoma. Also described are base sequences from the lp36
 CC position of human neuroblastoma cell lines (NB-1 and MASS-NB-SCH-1),
 CC which are tumour suppressor genes in human neuroblastoma. The genes are
 CC tumour markers and reagents in studying mechanism of tumour body
 CC formation, and gene diagnosis of tumors as well as in developing
 CC anti-cancer drugs. AAF9787 to AAF97829 represent PCR primers used in
 CC the exemplification of the present invention, and AAF97830 to AAF97874
 CC represent sequences given in the exemplification of the present
 XX SQ Sequence 34488 BP; 9654 A; 6717 C; 6926 G; 11191 T; 0 other;

Query Match 4.7%; Score 97.2; DB 22; Length 34488;
 Best Local Similarity 74.1%; Pred. No. 2.4e-12;
 Matches 123; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 258 AAATAGTATGACTCTGTTTCTTCAGCAGGACATATCTAAATAGGAGCTATACAAG 317
 Db 20646 AATAACTAGCAGCGTGCTCGCTTCGGCAGCATATATAAAATGGACGATACAG 20705

QY 318 AAGATTAGCATGACTCTGTGCAAGAAATGACACACAAATTTGTGAACATTCATATATT 377

Db 20706 AGATTAGCATGCGCCCTCGGCAAGATGACACGCAATTCGTGAACGTTCCATATT 20765

QY 378 AAAAAATAAATAATAAGAGAGAAAGGAAAAATTTAAAAAGAAA 423

KW anti allergic; anti asthmatic; anti inflammatory; olfactory;
XX respiratory active; ds.

OS Homo sapiens.

PN WO200155448-A1.

XX 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01333.

XX PF 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220363.

PR 26-JUL-2000; 2000US-0220364.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232197.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0244674.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.

PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.

PR 17-NOV-2000; 2000US-0249212.

PR 17-NOV-2000; 2000US-0249213.

PR 17-NOV-2000; 2000US-0249214.

PR 17-NOV-2000; 2000US-0249215.

PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.

PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.

PR 17-NOV-2000; 2000US-0249265.

PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.

PR 17-NOV-2000; 2000US-0249300.

PR 01-DEC-2000; 2000US-0250160.

PR 01-DEC-2000; 2000US-0250391.

PR 05-DEC-2000; 2000US-0251030.

PR 05-DEC-2000; 2000US-0251988.

PR 05-DEC-2000; 2000US-0256719.

PR 06-DEC-2000; 2000US-0251479.

PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251866.

PR 08-DEC-2000; 2000US-0251869.

PR 08-DEC-2000; 2000US-0251989.

PR 08-DEC-2000; 2000US-0251990.

PR 11-DEC-2000; 2000US-0254097.

PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-476224/51.
XX Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the respiratory system including respiratory
PT cancers and also for testing and detection e.g. diagnosis -
XX Disclosure; SEQ ID NO 1029; 546pp; English.
PS
XX The present invention relates to the isolation of novel human
CC respiratory antigens (AAU17685-AAU17975), and cDNA and genomic
CC sequences encoding for these polypeptides. The sequences of the
CC invention are useful for preventing, treating and/or prognosing
CC disorders related to the respiratory system including throat
CC disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),
CC lung disorders e.g. pneumonia, emphysema, nose disorders and cancers of
CC pleurisy, cystic fibrosis, allergic disorders e.g. asthma,
CC the respiratory tissues e.g. lung cancer. The polynucleotide sequences
CC of the invention are useful in gene therapy and antisense therapy.
CC AAS28161-AAS28764 represent genomic sequences encoding for novel
CC human respiratory antigens.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 6040 BP; 1446 A; 1488 C; 1420 G; 1686 T; 0 other;

Query Match 4.6%; Score 95.4; DB 22; Length 6040;
Best Local Similarity 72.8%; Pred. No. 3.5e-12;
Matches 123; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1136 ATATATAGTATTGTGGTAGAGATGGGATTTGCCATGTTGCCAGGCTAGTATGAAC 1195
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3417 ATTTTATATTTTATGTTGAGATGGGTTTCGCCATATTGGCAGGCTGTTGAAC 3358
QY 1196 CCTAAGCTAAGCAATCTCTGCTCTGCTCCCAAAATGTTGGGATTACAGGTGAAC 1255
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3357 CTGACCTAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3298
QY 1256 CACTGCACCGGCTGATAGTGTGTTTCATTACTTCTATTCTTGACCAC 1304
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3297 CACTGCACCGGCGGTTTTTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3249

RESULT 13
AAK84643/C
ID AAK84643 standard; DNA; 4181 BP.
XX
AC AAK84643;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:39455.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR

PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR

PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231124.
PR 08-SEP-2000; 2000US-02311243.
PR 08-SEP-2000; 2000US-02311244.
PR 08-SEP-2000; 2000US-0231143.
PR 08-SEP-2000; 2000US-0231144.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
(HUMA-) HUMAN GENOME SCI INC.
XX
Rosen CA, Barash SC, Ruben SM;
XX
WPI; 2001-483426/52.
XX
Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
Disclosure; SEQ ID NO 23094; 3071bp + Sequence Listing; English.
XX

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 04:10:57 ; Search time 608.653 Seconds
(without alignments)
10457.963 Million cell updates/sec

Title: US-10-005-337A-1
Perfect score: 2358
Sequence: 1 9gagcttcttgatttaacaa.....caggctcgaggccaccatgg 2358

Scoring table: OLIGO_NUC
Gap 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2467832

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database : N Geneseq 19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	33	1.4	48	AAH26804	Mouse orphan recep
2	21	0.9	22	AAZ37747	PCR primer AB3 for
3	21	0.9	22	AAZ50271	Primer AB3 to ampl
4	21	0.9	22	AAZ50660	Primer AB3 to ampl
5	21	0.9	22	AAH41892	Neomycin cassette
6	21	0.9	27	AAZ77092	PERK1 cloning prim
7	21	0.9	30	AAV10443	Plasmid araCB poly
8	21	0.9	30	AAA94281	Plasmid araCB poly

C 9	21	0.9	30	24	AAH26804	AraCB polylinker D
C 10	21	0.9	30	24	ABL57388	AraCB polylinker.
C 11	21	0.9	32	21	AAZ39184	Plasmid pUP1-1 con
C 12	21	0.9	33	17	AAZ27724	Primer B for amplfy
C 13	21	0.9	36	18	AAZ61299	Primer B for mouse
C 14	21	0.9	40	17	AAZ69476	Plasmid p1825f11 c
C 15	21	0.9	40	18	AAZ84401	M13mp18 single-str
C 16	21	0.9	40	18	AAZ846199	Capture DNA oligom
C 17	21	0.9	40	20	AAZ88900	Circular plasmid e
C 18	21	0.9	40	24	ABA98175	Oligonucleotide 40
C 19	21	0.9	50	16	AAQ96167	Stabiliser sequenc
C 20	21	0.9	50	18	AAZ47083	Synthetic target s
C 21	21	0.9	50	24	ABL51437	Bombay mori R2 ele
C 22	20	0.8	20	19	AAV67210	3' PCR primer P5 u
C 23	20	0.8	24	20	AAZ90613	FIV19 primer to cr
C 24	20	0.8	24	24	ABL60714	CMV promoter/enhan
C 25	20	0.8	25	19	AAV09826	Plasmid pTZ18R der
C 26	20	0.8	25	24	ABK89689	PCR primer #2 used
C 27	20	0.8	43	22	AAZ87604	Bacillus sp. KSM-A
C 28	19	0.8	19	24	AAZ44897	2038 PCR primer us
C 29	19	0.8	19	24	AAZ44911	DNA fragment II us
C 30	19	0.8	19	24	AAZ44914	PCR primer #2, to
C 31	19	0.8	20	16	AAQ75223	Amino labelled oli
C 32	19	0.8	20	16	AAQ75221	Amino labelled oli
C 33	19	0.8	20	24	ABK71151	Mouse HYPLIP1 locu
C 34	19	0.8	20	24	AAZ45762	Cancer cells detec
C 35	19	0.8	20	24	ABK68247	Mouse HYPLIP1 locu
C 36	19	0.8	22	16	AAQ95550	DNA probe 56 detec
C 37	19	0.8	24	24	AAZ45719	AML cancer cells d
C 38	19	0.8	24	24	AAZ45722	AML cancer cells d
C 39	19	0.8	25	24	AAZ33562	G6PDH-human CRP hy
C 40	19	0.8	38	24	AAZ44912	DNA fragment used
C 41	19	0.8	45	16	AAZ02881	Novel target nucle
C 42	19	0.8	50	16	AAZ02891	Novel target nucle
C 43	19	0.8	50	16	AAZ02880	Novel target nucle
C 44	19	0.8	50	16	AAZ02889	Novel target nucle
C 45	18	0.8	21	20	AAZ88524	Conus radiatus con

ALIGNMENTS

RESULT 1					
AAH26804					
ID	AAH26804	standard; DNA; 48 BP.			
XX					
AC	AAH26804;				
XX					
DT	26-NOV-2001	(first entry)			
XX					
DE	Mouse orphan receptor ROR2 gene-lacZ gene fusion construct.				
XX					
KW	Orphan receptor; ROR2; mouse; cartilage; chondrocyte;				
KW	osteoarthritis; therapy; mutant; beta-galactosidase; LacZ; ds.				
XX					
OS	Mus sp.				
OS	Synthetic.				
XX					
PH	Key	Location/Qualifiers			
FT	misc_feature	1..10			
FT		/tag= a			
FT		/note= "bases 1-10 are derived from the ROR2 gene"			
FT	misc_feature	11..45			
FT		/tag= b			
FT		/note= "bases 11-45 are derived from Bluescript polylinker"			
FT	misc_feature	46..48			
FT		/tag= c			
FT		/partial			
FT		/note= "lacZ coding region initiation codon"			
XX					
FN	WO200168812-A2.				
XX					

```

PD 20-SEP-2001.
XX
XX
XX 20-FEB-2001; 2001WO-US05473.
XX
XX 14-MAR-2000; 2000US-189216P.
XX
XX (REGE-) REGENERON PHARM INC.
XX
XX De Chiara TM, Kimble R, Yancopoulos GD;
XX
XX WPI; 2001-590052/66.
XX
XX P-PSDE; AAB82917.
XX
XX Increasing chondrocyte growth, development and activity, for increasing
XX cartilage formation, comprises contacting chondrocytes expressing
XX PT orphan receptor ROR2 with an agent capable of activating ROR2 receptor
XX
XX
XX Disclosure; Fig 1; 31pp; English.
XX
XX The present sequence is that of the fusion region between a mouse
XX ROR2 gene (see also AH26903) and the polylinker region of
XX Bluescript plasmid, which is flanked by a beta-galactosidase
XX (LacZ) coding region. Disruption of the endogenous murine ROR2
XX allele by insertion into Bluescript resulted in a sequence
XX encoding a novel product (see AAB82917) in which the ecto- and
XX transmembrane domains of ROR2 were fused to LacZ, which replaced
XX the normal tyrosine kinase-like cytodomain. This provided for
XX simultaneous knocking out of ROR2 function and the capability of
XX LacZ staining to follow ROR2 expression patterns. Mice heterozygous
XX for the mutant allele (ROR2+/-) were viable, fertile and appeared
XX normal. ROR2 was selectively expressed in the chondrocytes of the
XX developing anlagen of all bones formed by endochondral ossification.
XX Mice homozygous for the ROR2 mutant allele (ROR2-/-) exhibited
XX perinatal lethality, with shortened snout, limbs and tail as well
XX as cleft palate. Thus, disruption of the murine ROR2 gene leads to
XX profound skeletal abnormalities. As ROR2 appears to play a
XX critical role in cartilage formation it may be useful in developing
XX strategies to treat diseases of cartilage such as osteoarthritis.
XX Methods of increasing, or preventing, chondrocyte growth and
XX cartilage formation using agents that activate or block the ROR2
XX receptor are claimed, as well as methods of identifying agents
XX capable of activating, or blocking the activation, of the ROR2
XX receptor.
XX
XX Sequence 48 BP; 10 A; 13 C; 16 G; 9 T; 0 other;
SQ
Query Match 1.4%; Score 33; DB 22; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2325 AGCTTGCATGCTGCAGGTGGAGGCCACCATG 2357
Db 16 AGCTTGCATGCTGCAGGTGGAGGCCACCATG 48

RESULT 2
AAZ37747/c
ID AAZ37747 standard; DNA; 22 BP.
XX
XX AAZ37747;
XX
XX 01-FEB-2000 (first entry)
DT
XX
XX PCR primer AB3 for the isolation of BAC ends.
XX
XX Potato; Rx gene; resistance gene; potato virus X; PVX; transgenic plant;
XX broad spectrum extreme resistance; Narcissus mosaic virus; NMV; NVX; VMV;
XX Narcissus mosaic virus; Cymodinium mosaic virus; CymV; PopMV;
XX Poplar mosaic virus; White clover mosaic virus; WCLMV; PCR primer; ss.
XX
XX Synthetic.
XX
XX Solanum tuberosum.
OS

```

```

XX
XX WO954490-A2.
XX
XX 28-OCT-1999.
XX
XX 16-APR-1999; 99WO-GB01182.
XX
XX 16-APR-1998; 98GB-0008083.
XX
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX
XX Bendahmane A, Baulcombe DC, Kanyuka KV;
XX
XX WPI; 1999-634006/54.
XX
XX New isolated plant virus resistance gene, used to produce transgenic
XX plants with resistance to virus infection -
XX
XX Disclosure; Page 71; 124pp; English.
XX
XX PCR primers AAZ37745-237748 are used in the isolation and amplification
XX of the BAC end region in potato genomic DNA. The BAC sequences are used
XX to identify the potato Rx gene. The Rx gene (AAZ37153) is a resistance
XX gene which confers extreme resistance against potato virus X (PVX).
XX Sequence AAZ37154 is the Rx coding sequence free from introns. The Rx
XX gene can be used to create a recombinant vector which encodes the Rx
XX resistance polypeptide AAY52152, this vector can be used to transform
XX plant cells to produce a transgenic plant with resistance to PVX. The Rx
XX gene can be used to engineer resistance traits, preferably broad spectrum
XX extreme resistance, into plants. The Rx gene can also be activated by
XX non-PVX viruses, e.g. Narcissus mosaic virus (NMV), Nandina virus X
XX (NVX), Viola mosaic virus (VMV), Cymbidium mosaic virus (CYMV), Poplar
XX mosaic virus (PopMV) and White clover mosaic virus (WCLMV). Rx can be
XX used to offer specific protection against this group. The Rx gene
XX sequence can be used to create antibodies specific for Rx. The antibodies
XX can be used to down-regulate Rx activity and also for the detection,
XX identification or isolation of Rx or homologues.
XX
XX Sequence 22 BP; 5 A; 7 C; 6 G; 4 T; 0 other;
SQ
Query Match 0.9%; Score 21; DB 20; Length 22;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2325 AGCTTGCATGCTGCAGGTGG 2345
Db 21 AGCTTGCATGCTGCAGGTGG 1

RESULT 3
AAZ50271/c
ID AAZ50271 standard; DNA; 22 BP.
XX
XX AAZ50271;
XX
XX 18-MAY-2000 (first entry)
DT
XX
XX Primer AB3 to amplify right end sequence of the potato DNA insert.
XX
XX Resistance gene; Gpa2; potato; phytopathogenic nematode; chromosome 12p;
XX marker; IPMaC; 111R; Potato Cyst Nematode; PCN; Globodera; nematocide;
XX nucleotide binding site; NBS; nematode resistance; detection; diagnosis;
XX PCR primer; BAC clone; potato DNA insert; ss.
XX
XX Solanum tuberosum.
OS
XX
XX WO200006754-A2.
XX
XX 10-FEB-2000.
XX
XX 30-JUL-1999; 99WO-NL00491.
XX
XX 31-JUL-1998; 98WO-NL00445.
XX

```

XX (CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.
 PA (UYWA-) LANDBOOUNIVERSITEIT WAGENINGEN.
 XX
 PI Van Der Vossen EAG, Van Der Voort JNAMR, Lankhorst RMK, Bakker J;
 PI Stiekema WJ;
 XX
 DR WPI; 2000-183132/16.
 XX
 XX New isolated Gpa2 nematode resistance gene from potato, useful for
 PT producing plants which are resistant to nematodes and polypeptides for
 PT use in nematocidal compositions.
 XX
 PS Example 3; Page 26; 96pp; English.
 XX
 CC The present DNA sequence is the PCR primer AB3, used to amplify the
 CC right end sequence of the potato DNA insert in the BAC clone. The
 CC recircularised BAC clone with the potato DNA insert is used as the
 CC template for the amplification. This primer is used to screen Cara BAC
 CC library with markers IPW3, IPW4 and IPW5 and to isolate BAC clones from
 CC the Rxl/Gpa2 locus. The Gpa2 gene confers resistance to infections by
 CC phytopathogenic nematode of the Globodera genus. It is mapped to short
 CC arm of chromosome 12 (12p), between the markers IPW4c and 111R. This
 CC locus is associated with resistance to Potato Cyst Nematodes (PCN) like
 CC G. pallida and G. rostochiensis, that invade and damage the roots of
 CC Solanaceae. It has nematocidal activity. A recombinant DNA sequence
 CC comprising the Gpa2 gene can be transformed or transfected into
 CC plants, to provide increased resistance to nematodes. The polypeptides
 CC can also be used in nematocidal compositions and for detection and
 CC diagnosis of nematode infections.
 XX
 SQ Sequence 22 BP; 5 A; 7 C; 6 G; 4 T; 0 other;
 Query Match 0.9%; Score 21; DB 21; Length 22;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2325 AGCTTGCATGCTCGCAGGTGC 2345
 DB 21 AGCTTGCATGCTCGCAGGTGC 1
 RESULT 4
 AAZ50660/c
 ID AAZ50660 standard; DNA; 22 BP.
 XX
 AC AAZ50660;
 XX
 DT 23-MAY-2000 (first entry)
 DE
 DE Primer AB3 to amplify right end sequence of the potato DNA insert.
 XX
 XX Resistance; Gpa2; potato; phytopathogenic nematode; chromosome 12p;
 KW marker; IPW4c; 111R; Potato Cyst Nematode; PCN; Globodera; nematocidal;
 KW nematode resistance; detection; diagnosis; transgenic plant;
 KW PCR primer; ss.
 XX
 XX Solanum tuberosum.
 OS
 XX WO200006753-A1.
 PN
 XX 10-FEB-2000.
 PD
 XX 31-JUL-1998; 98WO-NL00445.
 PF
 XX 31-JUL-1998; 98WO-NL00445.
 PR
 XX (CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.
 PA (UYWA-) LANDBOOUNIVERSITEIT WAGENINGEN.
 XX
 PI Van Der Vossen EAG, Van Der Voort JNAMR, Lankhorst RMK, Bakker J;
 PI Stiekema WJ;
 XX

DR WPI; 2000-195310/17.
 XX
 PT Recombinant Gpa2 polynucleotide from potato, Solanum tuberosum, useful
 PT to confer resistance to phytopathic nematodes of the genus Globodera in
 PT transgenic plants, e.g. resistance to potato cyst nematodes in potatoes
 PT
 XX
 PS Example 3; Page 26; 96pp; English.
 XX
 CC The present sequence is the PCR primer AB3, used to amplify the
 CC right end sequence of the potato DNA insert in the BAC clone. The
 CC recircularised BAC clone with potato DNA insert is used as
 CC template for the amplification. This primer is used to screen Cara BAC
 CC library with markers IPW3, IPW4 and IPW5 and to isolate BAC clones from
 CC the Rxl/Gpa2 locus. Gpa2 gene confers resistance to infection by
 CC phytopathogenic nematode of the Globodera genus. It is mapped to short
 CC arm of chromosome 12 (12p), between the markers IPW4c and 111R. This
 CC locus is associated with resistance to Potato Cyst Nematodes (PCN) like
 CC G. pallida and G. rostochiensis, that invade and damage the roots of
 CC Solanaceae. It has nematocidal activity. Recombinant DNA sequence
 CC comprising the Gpa2 gene can be used to produce transgenic plants with
 CC increased resistance to nematodes. The polypeptides can also be used in
 CC nematocidal compositions and for detection and diagnosis of nematode
 CC infections.
 XX
 SQ Sequence 22 BP; 5 A; 7 C; 6 G; 4 T; 0 other;
 Query Match 0.9%; Score 21; DB 21; Length 22;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2325 AGCTTGCATGCTCGCAGGTGC 2345
 DB 21 AGCTTGCATGCTCGCAGGTGC 1
 RESULT 5
 AAH41892/c
 ID AAH41892 standard; DNA; 24 BP.
 XX
 AC AAH41892;
 XX
 DT 30-AUG-2001 (first entry)
 DE
 DE Neomycin cassette related PCR primer #4.
 XX
 KW Rapamycin; neomycin resistance; antifungal; anticancer; neurotrophin;
 KW immunosuppressant; polyketide synthase; cytostatic; neurotrophic;
 KW antitumour; PCR primer; ss.
 XX
 XX Streptomyces hygroscopicus.
 OS
 XX WO200134816-A1.
 PN
 XX 17-MAY-2001.
 PD
 XX 30-OCT-2000; 2000WO-US41701.
 PF
 XX 29-OCT-1999; 99US-0162229.
 PR
 XX 20-APR-2000; 2000US-0198578.
 XX
 XX (KOSA-) KOSAN BIOSCIENCES INC.
 XX
 XX Katz L, Liu L, Chung LM;
 PI
 DR WPI; 2001-355490/37.
 XX
 XX New recombinant Streptomyces hygroscopicus, for producing new rapamycin
 PT compounds, useful as e.g. anticancer agents, contains altered
 PT polyketide synthase genes -
 XX
 XX Example 2; Page 50; 62pp; English.
 XX

CC The present invention describes a recombinant Streptomyces hygroscopicus
CC (A) that produces rapamycin compounds (I). Also described are:
CC (1) substantially pure (I); (2) recombinant *S. hygroscopicus* that
CC expresses a hybrid PKS (polyketide synthase) comprising at least part
CC of each of rapamycin PKS and heterologous PKS; (3) recombinant
CC *S. hygroscopicus* that does not express at least one rapamycin modifying
CC enzyme but does produce a rapamycin analogue; and (4) recombinant
CC *S. hygroscopicus* that expresses a PKS consisting of only part of a
CC rapamycin PKS. (A) has cytosolic, antifungal, immunosuppressant and
CC neurotrophic activities. (I) are useful as antitumor, antifungal or
CC immunosuppressant agents and neurotrophics, and also as intermediates
CC for other active compounds. Modular construction of PKS makes possible
CC production of a large number of hybrids for screening. The present
CC sequence represents a PCR primer which is used in an example from the
CC present invention.

XX
SQ Sequence 24 BP; 5 A; 7 C; 7 G; 5 T; 0 other;
Query Match 0.9%; Score 21; DB 22; Length 24;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGC 2345
|||||
DB 23 AGCTTGCATGCTGCAGGTGC 3

RESULT 6
AAAF77092
ID AAF77092 standard; DNA; 27 BP.
XX
AC AAF77092;
XX
XX 17-MAY-2001 (first entry)
XX PERK1 cloning primer #1.
XX
XX Proline-rich extensin-like receptor kinase; PERK; resistance;
XX plant; ds.
XX
XX Brassica napus.
XX
XX WO200114563-A1.
XX
XX 01-MAR-2001.
XX
XX 18-AUG-2000; 2000WO-CA00966.
XX
XX 19-AUG-1999; 99US-0149466.
XX 13-OCT-1999; 99US-0159122.
XX
XX (GORI/) GORING D.
XX (SILV/) SILVA N.
XX
XX Goring D, Silva N;
XX
XX WPI; 2001-244305/25.
XX
XX New proline-rich, extensin-like receptor kinase nucleic acids and
XX polypeptides useful for increasing plant wounding or pathogen
XX resistance, or for producing transgenic plants with increased wounding
XX or pathogen resistance -
XX
XX Examples; Page 45; 91pp; English.

CC The present invention relates to proline-rich extensin-like
CC receptor kinase (PERK). The PERK nucleic acids and polypeptides
CC are useful for increasing the resistance of plants to wounding
CC and pathogens. These are also useful for producing transgenic
CC plants with increased wounding and pathogen resistance compared
CC with a wild type plant, as well as in assays for identifying
CC and developing compounds to inhibit and/or enhance polypeptide
CC function directly.

XX
SQ Sequence 27 BP; 6 A; 7 C; 9 G; 5 T; 0 other;
Query Match 0.9%; Score 21; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGC 2345
|||||
DB 5 AGCTTGCATGCTGCAGGTGC 25

RESULT 7
AAV10443/C
ID AAV10443 standard; cDNA; 30 BP.
XX
XX AAV10443;
XX
XX 15-JUN-1998 (first entry)
XX Plasmid araCB polylinker region.
XX
XX Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
XX Gaucher's disease; transgenic; detection; hybridisation; antifungal;
XX rheumatoid arthritis; overexpression; extracellular matrix;
XX arabinose; promoter; BAD operon; araC gene; ss.
XX
XX Synthetic.
XX
XX WO9747752-A1.
XX
XX 18-DEC-1997.
XX
XX 16-JUN-1997; 97WO-US10460.
XX
XX 14-JUN-1996; 96US-0663618.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gray PW;
XX
XX WPI; 1998-052316/05.
XX
XX Nucleic acids encoding human chitinase - useful as antifungal
XX agents, especially in combination with other antifungals
XX
XX Example 3; Page 15; 63pp; English.

CC AAV10441-V10443 are primers used to amplify the Salmonella typhimurium
CC arabinose promoter and araC gene. The resulting product is used in a
CC method of producing recombinant human chitinase in bacterial cells.
CC Chitinases are useful for treating or preventing fungal infection and
CC as immunogens for generating antibodies which are used to purify, detect
CC and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The
CC nucleic acid sequence of the chitinase is also useful as a probe to
CC identify and isolate genomic DNA encoding chitinases or similar
CC proteins, or cells expressing them or to generate transgenic ('knockout')
CC rodents. It can also be used in hybridisation assays and to detect
CC genetic alterations in the chitinase gene related to disease. Agents that
CC inhibit this protein may be useful in treatment of Gaucher's disease and
CC rheumatoid arthritis, where overexpression of the protein can damage the
CC extracellular matrix. Chitinase also improves the activity of other
CC antifungal agents and may allow a reduction in the dose of such agents,
CC and thus of their side effects.

XX
SQ Sequence 30 BP; 7 A; 8 C; 8 G; 7 T; 0 other;
Query Match 0.9%; Score 21; DB 19; Length 30;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGC 2345
|||||

```
Db      29 AGCTTGATGCTGCCTGCAGGTGC 9

RESULT 8
AAA94281/c
ID AAA94281 standard; DNA; 30 BP.
XX
XX
AC AAA94281;
XX
XX 08-JAN-2001 (first entry)
XX
XX Plasmid araCB polylinker region.
XX
XX araCB; arabinose operon; BAD; araC gene; calcineurin-binding peptide;
XX A-kinase anchor protein; AKAP; AKAP 79; immunostimulant;
XX interleukin 2 expression modulation; graft rejection; transplantation;
XX T cell-mediated disorder; ds.
XX
XX Synthetic.
XX
XX US6107104-A.
XX
XX 22-AUG-2000.
XX
XX 27-SEP-1996; 96US-0721458.
XX
XX 23-NOV-1994; 94US-0344227.
XX 15-MAR-1995; 95US-0404731.
XX 17-JUL-1995; 95US-0503226.
XX
XX (ICOS-) ICOS CORP.
XX
XX Lockerbie RO, Gallatin WM, Lai Y, Howard ML;
XX WPI; 2000-578541/54.
XX
XX Novel calcineurin deletion mutant having calcineurin polypeptide
XX sequence and binding A-kinase anchor proteins, for treating graft
XX rejection following organ transplantation and T cell-mediated disorders
XX
XX
XX Example 13; Column 23-24; 53pp; English.
XX
XX The present sequence is the polylinker region of plasmid araBC. araBC
XX contains the araC gene promoter, which was isolated from the arabinose
XX operon BAD of Salmonella typhimurium. It was used to express human
XX A-kinase anchor protein 79 (AKAP 79). By binding both cAMP-dependent
XX protein kinase (PKA) and calcineurin, AKAP 79 co-localises a kinase and a
XX phosphatase that may regulate flux through a specific signalling pathway.
XX Calcineurin is a Ca2+/calmodulin-dependent protein phosphatase which is
XX involved in many intracellular signalling pathways. It participates in
XX regulation of IL-2 expression following T cell stimulation in T cells.
XX Calcineurin-binding peptides derived from AKAP 79 may be used to inhibit
XX calcineurin activity in a cell. The peptides are useful for treating
XX graft rejection following organ transplantation and for treating
XX T cell-mediated disorders. Calcineurin deletion mutants which bind
XX AKAP 79 are useful for defining an AKAP 79 binding site, for stimulating
XX the immune response, stimulating activated T cells for selected clonal
XX expansion, or for enhancing T cell responses to experimental stimuli for
XX evaluation of early events in T cell biology and activation of the immune
XX response.
XX
XX Sequence 30 BP; 7 A; 8 C; 8 G; 7 T; 0 other;

Query Match      0.9%; Score 21; DB 21; Length 30;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2325 AGCTTGATGCTGCCTGCAGGTGC 2345
        |||||||
Db      29 AGCTTGATGCTGCCTGCAGGTGC 9

RESULT 9
AAD42624/c
ID AAD42624 standard; DNA; 30 BP.
XX
XX
AC AAD42624;
XX
XX 15-NOV-2002 (first entry)
XX
XX AraCB polylinker DNA.
XX
XX Chitinase; chitin-binding fragment; therapy; fungal infection;
XX fungicide; ss.
XX
XX Unidentified.
XX
XX US6399571-B1.
XX
XX 04-JUN-2002.
XX
XX 12-MAR-1999; 99US-0267574.
XX
XX 12-MAR-1998; 98US-0039198.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gray PW, Tjoelker LW;
XX WPI; 2002-606631/65.
XX
XX New chitinase derived polypeptides used for treating fungal infections
XX and for detecting and binding chitin
XX
XX Example 3; Column 43; 32pp; English.
XX
XX The invention relates to chitinase derived polypeptide, chitin-binding
XX fragments of human chitinase, fragment analogues, purified and isolated
XX polynucleotides encoding such fragments. The invention is used for
XX treating fungal infections and for detecting and binding chitin. The
XX present sequence is a araCB polylinker DNA used in the exemplification
XX of the invention.
XX
XX Sequence 30 BP; 7 A; 8 C; 8 G; 7 T; 0 other;

Query Match      0.9%; Score 21; DB 24; Length 30;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2325 AGCTTGATGCTGCCTGCAGGTGC 2345
        |||||||
Db      29 AGCTTGATGCTGCCTGCAGGTGC 9

RESULT 10
ABL57388/c
ID ABL57388 standard; DNA; 30 BP.
XX
XX
AC ABL57388;
XX
XX 12-AUG-2002 (first entry)
XX
XX AraCB polylinker.
XX
XX Arabinose; promoter; araC; chitinase; enzyme; fungicide;
XX antifungal; infection; ss.
XX
XX Salmonella typhimurium.
XX
XX US6372212-B1.
XX
XX 16-APR-2002.
XX
XX 16-JUN-1997; 97US-0877599.
XX
```

```

PR 14-JUN-1996; 96US-0663618.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gray PW;
XX
XX WPI; 2002-442449/47.
XX
XX Co-administering chitinase to improve the effectiveness of fungicidal
XX drugs e.g. amphotericin B or itraconazole, useful for treating fungal
XX infections e.g. candidiasis, coccidioidomycosis and blastomycosis -
XX
XX Example 3; Column 10; 26pp; English.
XX
XX The present sequence is that of an araCB polylinker flanked by a
XX 5' XbaI site and a 3' HindIII site. The polylinker was used in
XX the construction of expression plasmid pARMO218, in which
XX human chitinase cDNA (see ABL57380) was under the control of the
XX Salmonella arabinose promoter. The plasmid was used to transform
XX Escherichia coli cells for recombinant production of a chitinase
XX analogue (see ABB76234). The invention provides human chitinase
XX polynucleotides and polypeptides, and methods for the recombinant
XX production of human chitinase products, for use in treating fungal
XX infections. Human chitinase can be used to improve the activity of
XX a non-chitinase antifungal agent in the treatment of candidiasis,
XX aspergillosis, coccidioidomycosis, blastomycosis,
XX paracoccidioidomycosis, histoplasmosis, cryptococcosis,
XX chromoblastomycosis, sporotrichosis, mucormycosis, dermatophytoses
XX and Pneumocystis infections.
XX
XX Sequence 30 BP; 7 A; 8 C; 8 G; 7 T; 0 other;
SQ
Query Match 0.9%; Score 21; DB 24; Length 30;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2325 AGCTTGCATGCTCGAGGTCTG 2345
DB 29 AGCTTGCATGCTCGAGGTCTG 9
RESULT 11
AAZ39184/c
ID AAZ39184 standard; DNA; 32 BP.
XX
XX AAZ39184;
XX
XX 02-MAR-2000 (first entry)
XX
XX Plasmid pUP1-1 construction multiple cloning site oligonucleotide #2.
XX
XX Vaccine; delivery; hyperblebbing strain; Neisseria gonorrhoeae; bleb;
XX blebosome; antigen; diagnosis; immunisation; immunostimulant;
XX viral disease; rabies; hepatitis; bacterial disease; salmonella;
XX pneumonia; fungal disease; parasitic disease; infection; ss.
XX
XX Synthetic.
XX
XX WO9959625-A1.
XX
XX 25-NOV-1999.
XX
XX 10-MAY-1999; 99WO-US10033.
XX
XX 19-MAY-1998; 98US-0081576.
XX
XX (UYMA-) UNIV MARYLAND BALTIMORE.
XX
XX Stein DC;
XX
XX WPI; 2000-062379/05.
XX
XX Novel vaccine for providing immunity against disease -
PT

```

```

XX PS
XX XX
XX The present invention describes a vaccine for providing immunity against
XX a disease. The vaccine comprises blebsomes where an immunogenic
XX polypeptide specific for the disease is present in a pharmacologically
XX effective dose in a pharmaceutically acceptable excipient. The vaccine
XX is useful in the form of a pharmaceutical composition for treating a
XX disease, comprising a blebosome where the polypeptide (especially a
XX cytokine, a receptor and/or an antibiotic) is active against the disease.
XX The blebsomes are useful for the production of a vaccine for
XX immunisation against diseases and are useful in diagnostic assays where
XX the presence of antibodies against disease can be detected in samples
XX from a patient suspected of having the disease e.g. for the diagnosis of
XX viral diseases such as rabies or hepatitis, bacterial diseases such as
XX salmonella or pneumonia, fungal diseases and parasitic diseases. The
XX blebsomes are also useful in a delivery system for other biological
XX molecules e.g. chemotherapeutic agents for use in chemotherapy or immune
XX enhancers/suppressors. Vaccines using blebsomes collected from
XX hyperblebbing strains have an advantage over conventional vaccines as the
XX antigens are present in the absence of other cellular components. In
XX addition the antigens are assembled in a natural biological membrane
XX allowing the antigen to form a native conformation more closely mimicking
XX what is encountered in the natural organism. The present sequence
XX represents an oligonucleotide used in the exemplification of the present
XX invention.
XX
XX Sequence 32 BP; 7 A; 8 C; 10 G; 7 T; 0 other;
SQ
Query Match 0.9%; Score 21; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2325 AGCTTGCATGCTCGAGGTCTG 2345
DB 29 AGCTTGCATGCTCGAGGTCTG 9
RESULT 12
AAT27724/c
ID AAT27724 standard; DNA; 33 BP.
XX
XX AAT27724;
XX
XX 08-NOV-1996 (first entry)
XX
XX Primer for amplifying S. aureus toxic shock syndrome toxin gene.
XX
XX Toxic shock syndrome; toxin; TSST-1; Staphylococcus aureus;
XX superantigen; mutant; derivative; deletion; substitution; vaccine;
XX treatment; prophylactic; antibody response; T-cell; thymocyte; MHC;
XX major histocompatibility complex; ss.
XX
XX Synthetic.
XX
XX WO9614744-A1.
XX
XX 23-MAY-1996.
XX
XX 08-NOV-1995; 95WO-US14639.
XX
XX 14-NOV-1994; 94US-0338373.
XX
XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
XX Kappler JW, Marrack P, Matsumura M, Shimonkevitz R;
XX
XX WPI; 1996-259474/26.
XX
XX Modification of T-cell responsiveness using S. aureus TSST-1
XX super-antigen or deriv. - useful for preventing and treating
XX antigen-mediated and antigen-initiated diseases, such as toxic shock
XX syndrome
PT

```


XX Example 2; Page 42; 71pp; English.

XX The Staphylococcus aureus toxic shock syndrome toxin-1 (TSST-1) is

CC modified or mutated so that it no longer has the pathological effect

CC of a superantigen but is still capable of eliciting an antibody

CC response which cross reacts with and protects against the native

CC superantigen. The TSST-1 superantigen derivative modifies T-cell

CC receptors or class II major histocompatibility complex (MHC) binding.

CC The superantigen derivative is preferably a deletion or substitution

CC mutation of a portion of the superantigen. The derivatives are

CC useful as a vaccine for preventing and treating antigen-mediated and

CC antigen-initiated diseases. Two primers (AAT27723, AAT27724) were used

CC in the amplification, subcloning and mutagenesis of the toxic shock

CC syndrome toxin gene.

XX Sequence 33 BP; 7 A; 9 C; 9 G; 8 T; 0 other;

XX

Query Match 0.9%; Score 21; DB 17; Length 33;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGC 2345

Db 32 AGCTTGCATGCTGCAGGTGC 12

RESULT 13

AAT61299

ID AAT61299 standard; DNA; 36 BP.

XX AAT61299;

DT 12-NOV-1997 (first entry)

XX Primer B for mouse T cell receptor zeta chain.

XX

XX Primer; polymerase chain reaction; PCR; amplification; mouse;

KW murine; extracellular domain; major histocompatibility complex;

KW MHC; class I; signal transduction; molecule; T cell; receptor;

KW zeta chain; B cell; killer cell; prevention; transplant;

KW xenograft; allograft; rejection; treatment; autoimmune disease;

KW systemic lupus erythematosus; myasthenia gravis;

KW rheumatoid arthritis; insulin dependent; diabetes mellitus; ss.

XX

OS Synthetic.

XX

XX WO9708328-A1.

PN 06-MAR-1997.

XX

XX 29-AUG-1996; 96WO-US13873.

XX

XX 30-AUG-1995; 95US-0002964.

XX

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA

XX Rosenberg AS;

PI WPI; 1997-179285/16.

XX

XX Inhibiting T cells that recognise preselected target molecules -

PT using killer cells having signal transduction molecule fused to

PT target molecule, useful for inhibiting graft rejection or treating

PT auto-immune disease

XX

XX Example 1; Page 38; 157pp; English.

XX

XX The present sequence is a primer for the PCR amplification of a

CC cDNA encoding the mouse T cell receptor zeta chain. A novel cell,

CC comprising a signal transduction molecule, preferably a mouse

CC T cell receptor zeta chain, attached to the extracellular domain

CC (5D) of a major histocompatibility complex (MHC) molecule,

CC preferably the ED of a mouse MHC class I molecule, can be used to

CC improve T cell activation, when it is a B cell, or prolong the life

CC of tissue cells in a host having a T cell that specifically

CC recognises and mediates an immune response directed against the

CC tissue, when it is a killer cell. Specifically, this is useful for

CC preventing transplant (xenograft or allograft) rejection and

CC treating autoimmune disease, e.g. systemic lupus erythematosus,

CC myasthenia gravis, rheumatoid arthritis and insulin dependent

CC diabetes mellitus.

XX

XX Sequence 36 BP; 8 A; 10 C; 9 G; 9 T; 0 other;

XX

Query Match 0.9%; Score 21; DB 18; Length 36;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGC 2345

Db 14 AGCTTGCATGCTGCAGGTGC 34

RESULT 14

AAT69476

ID AAT69476 standard; DNA; 40 BP.

XX AAT69476;

AC

XX 29-JUL-1997 (first entry)

DT

XX Plasmid p182SfII component oligonucleotide 40R.

DE

XX

XX Oligonucleotide; plasmid; p182SfII; gene; promoter; R-TEM1;

KW beta-lactamase; construction; assembly; synthesis; end;

KW complementary; polymerase; reaction; ECPR; ss.

XX

OS Synthetic.

XX

XX WO9633207-A1.

PN 24-OCT-1996.

XX

XX 18-APR-1996; 96WO-US05480.

XX

XX 18-APR-1995; 95US-0425684.

XX

XX (GLAX) GLAXO GROUP LTD.

PA

XX Lipshutz RJ, Stemmer WPC;

PI WPI; 1996-485725/48.

XX

XX Polynucleotide amplification using bivalent PCR primer to generate

PT rolling circle intermediate or linear concatamers - requires reduced

PT amt. of primer compared to conventional PCR

XX

XX Example 2; Fig 5C; 77pp; English.

XX

XX The present sequence is a component oligonucleotide of the 2.71 kb

CC plasmid p182SfII (Stemmer (1994) Nature 370:389), which encodes the

CC gene and promoter region for R-TEM1 beta-lactamase. A collection of

CC 134 component oligonucleotides, including the present sequence,

CC were synthesised and used to assemble the plasmid by end

CC complementary polymerase reaction (ECPR), employing the overlapping

CC ends of the oligonucleotides.

XX

XX Sequence 40 BP; 7 A; 13 C; 12 G; 8 T; 0 other;

XX

Query Match 0.9%; Score 21; DB 17; Length 40;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGC 2345

```
Db      15 AGCTTGCATGCTGCAGGTGC 35

RESULT 15
AAT88401/c
ID      AAT88401 standard; DNA; 40 BP.
XX
AC      AAT88401;
XX
XX      02-APR-1998 (first entry)
XX
DE      M13mp18 single-stranded DNA probe.
XX
KW      Probe; detection; RecA protein; ligase; circular DNA probe; label;
KW      ss.
XX
OS      Synthetic.
XX
FH      Key                      Location/Qualifiers
FT      modified_base 1
FT      /*tag= a
FT      /note= "Cytosine modified by a label"
XX
XX      JP09220099-A.
XX
XX      26-AUG-1997.
XX
XX      15-FEB-1996; 96JP-0054278.
XX
XX      15-FEB-1996; 96JP-0054278.
XX
XX      (AISH-) AISHIN COSMOS KENKYUSHO KK.
XX
XX      WPI; 1997-474312/44.
XX
PT      Using single stranded DNA probe to detect double stranded DNA
PT      sequence - having specific nucleotide sequence
XX
XX      Disclosure; Page 5; 7pp; Japanese.
XX
CC      The present sequence represents an M13mp18 single stranded DNA probe
CC      which was used in a new method for detecting double stranded DNA having
CC      a specific DNA nucleotide sequence. The method comprises: (a) mixing the
CC      single stranded probe, labelled at the 5' end and with specific
CC      nucleotide sequences at its 3' and 5' ends, with the RecA protein
CC      molecule; (b) adding a DNA sample containing the double stranded DNA
CC      sequence, which is complementary to the specific nucleotide sequences;
CC      (c) adding ligase, allowing the nucleotide sequence of the double
CC      stranded DNA to overlap with the specific nucleotide sequence at the ends
CC      of the probe and allowing the ends bind each other to give a circular DNA
CC      probe; and (d) removing the RecA protein molecule, and detecting the
CC      resulting hybrid DNA using the label. The method uses to a single
CC      stranded DNA probe to detect a double stranded DNA having a specific
CC      nucleotide sequence. A nucleic acid sequence can be detected without
CC      damage, allowing accurate sequence information to be obtained.
XX
SQ      Sequence 40 BP; 8 A; 12 C; 13 G; 7 T; 0 other;

Query Match      0.9%; Score 21; DB 18; Length 40;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2325 AGCTTGCATGCTGCAGGTGC 2345
          |||||
Db      28 AGCTTGCATGCTGCAGGTGC 8

Search completed: November 14, 2003, 11:58:53
Job time : 610.853 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 11:17:54 ; Search time 4903.81 Seconds
(without alignments)
11686.822 Million cell updates/sec

Title: US-10-005-337A-1
Perfect score: 2358
Sequence: 1 ggaatcccttcattggttaaca.....caggctggaggccaccatgg 2358

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0
Total number of hits satisfying chosen parameters: 124404

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pin.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	21	0.9	31	29	DR19H5T
C 2	21	0.9	31	29	DR6J9T
C 3	19	0.8	30	29	DR43H24T
C 4	19	0.8	30	29	DR43J2T

C 5	19	0.8	31	29	DR43B24T	AL986461	Danio rer
C 6	19	0.8	31	29	DR43B2T	AL986536	Danio rer
C 7	19	0.8	31	29	DR48K10T	AL978349	Danio rer
C 8	19	0.8	32	29	DR1009T	AL745332	Danio rer
C 9	18	0.8	30	29	DR42M15T	AL978871	Danio rer
C 10	18	0.8	32	29	DR1N11T	AL736520	Danio rer
C 11	18	0.8	33	29	DR43F16T	AL975436	Danio rer
C 12	18	0.8	35	29	DR10P1T	AL975949	Danio rer
C 13	18	0.8	36	29	DR43A15T	AL977394	Danio rer
C 14	18	0.8	37	29	DR43P8T	AL977007	Danio rer
C 15	18	0.8	43	29	DR35H17T	AL975957	Danio rer
C 16	18	0.8	50	12	BM253047	BM253047	Kf01_K Pl
C 17	17	0.7	22	28	AZ346808	AZ346808	1M0082P12
C 18	17	0.7	30	29	DR43K4T	AL984225	Danio rer
C 19	17	0.7	31	29	DR43O8T	AL985946	Danio rer
C 20	17	0.7	31	29	DR48D10T	AL979088	Danio rer
C 21	17	0.7	32	29	DR43I18T	AL984188	Danio rer
C 22	17	0.7	33	29	DR43C9T	AL981006	Danio rer
C 23	17	0.7	33	29	DR43K14T	AL978734	Danio rer
C 24	17	0.7	34	29	DR43K2T	AL975040	Danio rer
C 25	17	0.7	36	29	DR43H15T	AL981997	Danio rer
C 26	17	0.7	36	29	DR43L19T	AL987312	Danio rer
C 27	17	0.7	36	29	DR43M4T	AL981768	Danio rer
C 28	17	0.7	38	29	DR43C7T	AL979995	Danio rer
C 29	17	0.7	39	29	DR1L3T	AL975919	Danio rer
C 30	17	0.7	39	29	DR43P11T	AL987184	Danio rer
C 31	16	0.7	28	14	T81587	T81587	Yd44C02.81
C 32	16	0.7	30	29	DR43D1T	AL975931	Danio rer
C 33	16	0.7	31	29	DR43J7T	AL984342	Danio rer
C 34	16	0.7	34	29	DR49M12T	AL986978	Danio rer
C 35	16	0.7	37	28	AZ819838	AZ819838	2M0091M11
C 36	16	0.7	37	29	DR43D23T	AL984974	Danio rer
C 37	16	0.7	37	29	DR43N7T	AL975774	Danio rer
C 38	16	0.7	38	9	AU008661	AU008661	AU008661
C 39	16	0.7	38	29	DR43L9T	AL973965	Danio rer
C 40	16	0.7	38	29	DR48J2T	AL982936	Danio rer
C 41	16	0.7	48	9	AU257958	AU257958	AU257958
C 42	16	0.7	48	9	AU263470	AU263470	AU263470
C 43	15	0.6	30	29	DR7D7T	AL745225	Danio rer
C 44	15	0.6	31	29	DR2023T	AL743836	Danio rer
C 45	15	0.6	32	29	DR1L19T	AL741484	Danio rer

ALIGNMENTS

RESULT 1
DR19H5T/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
Source

DR19H5T
Danio rerio genomic clone DKEY-19H5, genomic survey sequence.
AL735758
GSS: AL735758.1 GI:21344389
Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 31)
Humphray,S.J., Huckle,E. and Hunt,S.E.
Direct Submission
Submitted (06-JUN-2002) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact: humquerry@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 19H5. 19H5 is part of the Daniokey BAC Library created by R. Plasterk and N.V. Keygene.
Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

```

/clone="DKEY-19H5"
/tissue_type="Testis"
/note="vector pIndigoBAC-536"
8 a 10 c 8 g 5 t

Query Match 0.9%; Score 21; DB 29; Length 31;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGC 2345
Db 31 AGCTTGCATGCTGCAGGTGC 11

RESULT 2
LOCUS DR6J9T/c 31 bp DNA linear GSS 27-NOV-2002
DEFINITION Danio rerio genomic clone DKEY-6J9, genomic survey sequence.
ACCESSION AL748020
VERSION AL748020.1 GI:21353176
KEYWORDS GSS.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE
AUTHORS Humphray, S.J., Huckle, E. and Hunt, S.E.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2002) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
humquery@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 6J9. 6J9 is part
of the Danokey BAC Library created by R. Plasterk and N.V.
Keygene.
Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES
Location/Qualifiers
1..31
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-6J9"
/tissue_type="Testis"
/note="vector pIndigoBAC-536"
8 a 10 c 8 g 5 t

BASE COUNT 8 a 10 c 8 g 5 t
ORIGIN

Query Match 0.9%; Score 21; DB 29; Length 31;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGC 2345
Db 31 AGCTTGCATGCTGCAGGTGC 11

RESULT 3
LOCUS DR43H24T/c 30 bp DNA linear GSS 22-NOV-2002
DEFINITION Danio rerio genomic clone DKEY-43H24, genomic survey sequence.
ACCESSION AL983803
VERSION AL983803.1 GI:25184459
KEYWORDS GSS.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE
AUTHORS Humphray, S.J., Huckle, E. and Hunt, S.E.
TITLE Direct Submission
JOURNAL Submitted (14-NOV-2002) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
humquery@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 43J2. 43J2 is
part of the Danokey BAC Library created by R. Plasterk and N.V.
Keygene.
Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES
Location/Qualifiers
1..30
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-43J2"
/tissue_type="Testis"
/note="vector pIndigoBAC-536"
7 a 9 c 7 g 7 t

BASE COUNT 7 a 9 c 7 g 7 t
ORIGIN

Query Match 0.8%; Score 19; DB 29; Length 30;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGT 2343
Db 30 AGCTTGCATGCTGCAGGT 12

RESULT 4
LOCUS DR43J2T/c 30 bp DNA linear GSS 22-NOV-2002
DEFINITION Danio rerio genomic clone DKEY-43J2, genomic survey sequence.
ACCESSION AL984851
VERSION AL984851.1 GI:25184200
KEYWORDS GSS.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE
AUTHORS Humphray, S.J., Huckle, E. and Hunt, S.E.
TITLE Direct Submission
JOURNAL Submitted (14-NOV-2002) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
humquery@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 43J2. 43J2 is
part of the Danokey BAC Library created by R. Plasterk and N.V.
Keygene.
Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES
Location/Qualifiers
1..30
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-43J2"
/tissue_type="Testis"
/note="vector pIndigoBAC-536"
7 a 9 c 7 g 7 t

BASE COUNT 7 a 9 c 7 g 7 t
ORIGIN

Query Match 0.8%; Score 19; DB 29; Length 30;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGT 2343
Db 30 AGCTTGCATGCTGCAGGT 12

RESULT 5
LOCUS DR43B24T/c 31 bp DNA linear GSS 22-NOV-2002
DEFINITION Danio rerio genomic clone DKEY-43B24, genomic survey sequence.
ACCESSION AL986461

```

```

Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
humquery@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 43H24. 43H24 is
part of the Danokey BAC Library created by R. Plasterk and N.V.
Keygene.
Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES
Location/Qualifiers
1..30
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-43H24"
/tissue_type="Testis"
/note="vector pIndigoBAC-536"
7 a 9 c 7 g 7 t

BASE COUNT 7 a 9 c 7 g 7 t
ORIGIN

Query Match 0.8%; Score 19; DB 29; Length 30;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGT 2343
Db 30 AGCTTGCATGCTGCAGGT 12

RESULT 5
LOCUS DR43B24T/c 31 bp DNA linear GSS 22-NOV-2002
DEFINITION Danio rerio genomic clone DKEY-43B24, genomic survey sequence.
ACCESSION AL986461

```

```

VERSION      AL986461.1  GI:25184019
KEYWORDS     GSS.
SOURCE       Danio rerio (zebrafish)
ORGANISM     Danio rerio

REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
TITLE        Cypriniformes; Cyprinidae; Danio.
JOURNAL      1 (bases 1 to 31)
             Humphray,S.J., Huckle,E. and Hunt,S.E.
             Direct Submission
             Submitted (14-NOV-2002) The Sanger Institute, Wellcome Trust Genome
             Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
             humquery@sanger.ac.uk Unpublished
COMMENT      This sequence was generated from the T7 end of BAC 43B24. 43B24 is
             part of the Daniokey BAC Library created by R. Plasterk and N.V.
             Keygene.
             Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES     Location/Qualifiers
             1..31
             /organism="Danio rerio"
             /mol_type="genomic DNA"
             /db_xref="taxon:7955"
             /clone="DKEY-43B24"
             /tissue_type="Testis"
             /note="vector pindigoBAC-536"
BASE COUNT   7 a 9 c 7 g 8 t
ORIGIN
Query Match      0.8%; Score 19; DB 29; Length 31;
Best Local Similarity 100.0%; Pred.No. 7.8e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCGCTGCAGGT 2343
Db 30 AGCTTGCATGCGCTGCAGGT 12

RESULT 6
DR43B2T/c
LOCUS
DEFINITION    Danio rerio genomic clone DKEY-43B2, genomic survey sequence.
ACCESSION     AL986536
VERSION       AL986536.1  GI:25183707
KEYWORDS      GSS.
SOURCE        Danio rerio (zebrafish)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
              Cypriniformes; Cyprinidae; Danio.
              Humphray,S.J., Huckle,E. and Hunt,S.E.
              Direct Submission
              Submitted (14-NOV-2002) The Sanger Institute, Wellcome Trust Genome
              Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
              humquery@sanger.ac.uk Unpublished
COMMENT      This sequence was generated from the T7 end of BAC 43B2. 43B2 is
              part of the Daniokey BAC Library created by R. Plasterk and N.V.
              Keygene.
              Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES     Location/Qualifiers
             1..31
             /organism="Danio rerio"
             /mol_type="genomic DNA"
             /db_xref="taxon:7955"
             /clone="DKEY-43B2"
             /tissue_type="Testis"
             /note="vector pindigoBAC-536"
BASE COUNT   7 a 9 c 7 g 8 t
ORIGIN
Query Match      0.8%; Score 19; DB 29; Length 31;
Best Local Similarity 100.0%; Pred.No. 7.8e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCGCTGCAGGT 2343
Db 30 AGCTTGCATGCGCTGCAGGT 12

RESULT 6
DR43B2T/c
LOCUS
DEFINITION    Danio rerio genomic clone DKEY-43B2, genomic survey sequence.
ACCESSION     AL986536
VERSION       AL986536.1  GI:25183707
KEYWORDS      GSS.
SOURCE        Danio rerio (zebrafish)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
              Cypriniformes; Cyprinidae; Danio.
              Humphray,S.J., Huckle,E. and Hunt,S.E.
              Direct Submission
              Submitted (14-NOV-2002) The Sanger Institute, Wellcome Trust Genome
              Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
              humquery@sanger.ac.uk Unpublished
COMMENT      This sequence was generated from the T7 end of BAC 43B2. 43B2 is
              part of the Daniokey BAC Library created by R. Plasterk and N.V.
              Keygene.
              Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES     Location/Qualifiers
             1..31
             /organism="Danio rerio"
             /mol_type="genomic DNA"
             /db_xref="taxon:7955"
             /clone="DKEY-43B2"
             /tissue_type="Testis"
             /note="vector pindigoBAC-536"
BASE COUNT   7 a 9 c 7 g 8 t
ORIGIN
Query Match      0.8%; Score 19; DB 29; Length 31;
Best Local Similarity 100.0%; Pred.No. 7.8e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCGCTGCAGGT 2343
Db 30 AGCTTGCATGCGCTGCAGGT 12

```

```

QY 2325 AGCTTGCATGCGCTGCAGGT 2343
Db 30 AGCTTGCATGCGCTGCAGGT 12

RESULT 7
DR48K10T/c
LOCUS
DEFINITION    Danio rerio genomic clone DKEY-48K10, genomic survey sequence.
ACCESSION     AL978349
VERSION       AL978349.1  GI:25186345
KEYWORDS      GSS.
SOURCE        Danio rerio (zebrafish)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
              Cypriniformes; Cyprinidae; Danio.
              Humphray,S.J., Huckle,E. and Hunt,S.E.
              Direct Submission
              Submitted (14-NOV-2002) The Sanger Institute, Wellcome Trust Genome
              Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
              humquery@sanger.ac.uk Unpublished
COMMENT      This sequence was generated from the T7 end of BAC 48K10. 48K10 is
              part of the Daniokey BAC Library created by R. Plasterk and N.V.
              Keygene.
              Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES     Location/Qualifiers
             1..31
             /organism="Danio rerio"
             /mol_type="genomic DNA"
             /db_xref="taxon:7955"
             /clone="DKEY-48K10"
             /tissue_type="Testis"
             /note="vector pindigoBAC-536"
BASE COUNT   7 a 9 c 7 g 8 t
ORIGIN
Query Match      0.8%; Score 19; DB 29; Length 31;
Best Local Similarity 100.0%; Pred.No. 7.8e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCGCTGCAGGT 2343
Db 30 AGCTTGCATGCGCTGCAGGT 12

RESULT 8
DR1009T/c
LOCUS
DEFINITION    Danio rerio genomic clone DKEY-1009, genomic survey sequence.
ACCESSION     AL745332
VERSION       AL745332.1  GI:21339021
KEYWORDS      GSS.
SOURCE        Danio rerio (zebrafish)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
              Cypriniformes; Cyprinidae; Danio.
              Humphray,S.J., Huckle,E. and Hunt,S.E.
              Direct Submission
              Submitted (06-JUN-2002) The Sanger Institute, Wellcome Trust Genome
              Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
              humquery@sanger.ac.uk Unpublished
COMMENT      This sequence was generated from the T7 end of BAC 1009. 1009 is
              part of the Daniokey BAC Library created by R. Plasterk and N.V.
              Keygene.
              Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES     Location/Qualifiers
             1..32
             /organism="Danio rerio"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-1009"
/tissue_type="Testis"
/note="vector pindigobAC-536"
7 g
7 t

BASE COUNT      8 a      10 c      7 g
ORIGIN

Query Match      0.8%; Score 19; DB 29; Length 32;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGT 2342
|||||
Db 31 AGCTTGCATGCTGCAGGT 13

RESULT 9
DR42M15T/c
LOCUS
DEFINITION
Danio rerio genomic clone DKEY-42M15, genomic survey sequence.
ACCESSION
AL978871
VERSION
AL978871.1 GI:25183839
KEYWORDS
GSS.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 30)
Humphray,S.J., Huckle,E. and Hunt,S.E.
Direct Submission
Submitted (14-NOV-2002) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
humquerry@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 42M15 is
part of the Daniokey BAC Library created by R. Plasterk and N.V.
Keygene.
Further details: http://www.sanger.ac.uk/Projects/D_rerio/.
Location/Qualifiers
1..30
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-42M15"
/tissue_type="Testis"
/note="vector pindigobAC-536"
7 g
8 t

BASE COUNT      6 a      9 c      7 g
ORIGIN

Query Match      0.8%; Score 18; DB 29; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.4e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGG 2342
|||||
Db 29 AGCTTGCATGCTGCAGG 12

RESULT 10
DR1N11T/c
LOCUS
DEFINITION
Danio rerio genomic clone DKEY-IN11, genomic survey sequence.
ACCESSION
AL736520
VERSION
AL736520.1 GI:21344444
KEYWORDS
GSS.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 32)
Humphray,S.J., Huckle,E. and Hunt,S.E.
Direct Submission
Submitted (14-NOV-2002) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
humquerry@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 43F16. 43F16 is
part of the Daniokey BAC Library created by R. Plasterk and N.V.
Keygene.
Further details: http://www.sanger.ac.uk/Projects/D_rerio/.
Location/Qualifiers
1..33
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-43F16"
/tissue_type="Testis"
/note="vector pindigobAC-536"
8 g
10 t

BASE COUNT      6 a      9 c      8 g
ORIGIN

Query Match      0.8%; Score 18; DB 29; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGG 2342
|||||
Db 30 AGCTTGCATGCTGCAGG 13

RESULT 12
DR10P1T/c
LOCUS

```

```

TITLE
JOURNAL
Direct Submission
Submitted (06-JUN-2002) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
humquerry@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC IN11. IN11 is
part of the Daniokey BAC Library created by R. Plasterk and N.V.
Keygene.
Further details: http://www.sanger.ac.uk/Projects/D_rerio/.
Location/Qualifiers
1..32
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-IN11"
/tissue_type="Testis"
/note="vector pindigobAC-536"
7 a      11 c      8 g      6 t

BASE COUNT      7 a      11 c      8 g
ORIGIN

Query Match      0.8%; Score 18; DB 29; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGG 2342
|||||
Db 32 AGCTTGCATGCTGCAGG 15

RESULT 11
DR43F16T/c
LOCUS
DEFINITION
Danio rerio genomic clone DKEY-43F16, genomic survey sequence.
ACCESSION
AL975436
VERSION
AL975436.1 GI:25184284
KEYWORDS
GSS.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 33)
Humphray,S.J., Huckle,E. and Hunt,S.E.
Direct Submission
Submitted (14-NOV-2002) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
humquerry@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 43F16. 43F16 is
part of the Daniokey BAC Library created by R. Plasterk and N.V.
Keygene.
Further details: http://www.sanger.ac.uk/Projects/D_rerio/.
Location/Qualifiers
1..33
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-43F16"
/tissue_type="Testis"
/note="vector pindigobAC-536"
8 g
10 t

BASE COUNT      6 a      9 c      8 g
ORIGIN

Query Match      0.8%; Score 18; DB 29; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGG 2342
|||||
Db 30 AGCTTGCATGCTGCAGG 13

RESULT 12
DR10P1T/c
LOCUS

```

```

DEFINITION      Danio rerio genomic clone DKEY-10P1, genomic survey sequence.
ACCESSION       AL737949
VERSION         AL737949.1  GI:21338585
KEYWORDS        GSS.
SOURCE          Danio rerio (zebrafish)
ORGANISM        Danio rerio
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
                Cypriniformes; Cyprinidae; Danio.
REFERENCE       1 (bases 1 to 35)
AUTHORS         Humphray,S.J., Huckle,E. and Hunt,S.E.
TITLE           Direct Submission
JOURNAL         Submitted (06-JUN-2002) The Sanger Institute, Wellcome Trust Genome
                Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
                humquerry@sanger.ac.uk Unpublished
COMMENT         This sequence was generated from the T7 end of BAC 10P1. 10P1 is
                part of the Daniokey BAC Library created by R. Plasterk and N.V.
                Keygene.
                Further details: http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES        Location/Qualifiers
                 1..35
                    /organism="Danio rerio"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:7955"
                    /clone="DKEY-10P1"
                    /tissue_type="Testis"
                    /notes="vector pindigobAC-536"
BASE COUNT      9 a 11 c 8 g 7 t
ORIGIN           9 a 11 c 8 g 7 t

Query Match      0.8%; Score 18; DB 29; Length 35;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCGCTGCAGG 2342
Db 33 AGCTTGCATGCGCTGCAGG 16

RESULT 13
LOCUS            DR43A15T/c
DEFINITION      Danio rerio genomic clone DKEY-43A15, genomic survey sequence.
ACCESSION       AL973794
VERSION         AL973794.1  GI:25183730
KEYWORDS        GSS.
SOURCE          Danio rerio (zebrafish)
ORGANISM        Danio rerio
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
                Cypriniformes; Cyprinidae; Danio.
REFERENCE       1 (bases 1 to 36)
AUTHORS         Humphray,S.J., Huckle,E. and Hunt,S.E.
TITLE           Direct Submission
JOURNAL         Submitted (14-NOV-2002) The Sanger Institute, Wellcome Trust Genome
                Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
                humquerry@sanger.ac.uk Unpublished
COMMENT         This sequence was generated from the T7 end of BAC 43A15. 43A15 is
                part of the Daniokey BAC Library created by R. Plasterk and N.V.
                Keygene.
                Further details: http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES        Location/Qualifiers
                 1..36
                    /organism="Danio rerio"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:7955"
                    /clone="DKEY-43A15"
                    /tissue_type="Testis"
                    /notes="vector pindigobAC-536"
BASE COUNT      7 a 9 c 9 g 11 t
ORIGIN           7 a 9 c 9 g 11 t

Query Match      0.8%; Score 18; DB 29; Length 36;

```

```

Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCGCTGCAGG 2342
Db 30 AGCTTGCATGCGCTGCAGG 13

RESULT 14
LOCUS            DR43P8T/c
DEFINITION      Danio rerio genomic clone DKEY-43P8, genomic survey sequence.
ACCESSION       AL977007
VERSION         AL977007.1  GI:25184098
KEYWORDS        GSS.
SOURCE          Danio rerio (zebrafish)
ORGANISM        Danio rerio
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
                Cypriniformes; Cyprinidae; Danio.
REFERENCE       1 (bases 1 to 37)
AUTHORS         Humphray,S.J., Huckle,E. and Hunt,S.E.
TITLE           Direct Submission
JOURNAL         Submitted (14-NOV-2002) The Sanger Institute, Wellcome Trust Genome
                Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
                humquerry@sanger.ac.uk Unpublished
COMMENT         This sequence was generated from the T7 end of BAC 43P8. 43P8 is
                part of the Daniokey BAC Library created by R. Plasterk and N.V.
                Keygene.
                Further details: http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES        Location/Qualifiers
                 1..37
                    /organism="Danio rerio"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:7955"
                    /clone="DKEY-43P8"
                    /tissue_type="Testis"
                    /notes="vector pindigobAC-536"
BASE COUNT      7 a 10 c 11 g 9 t
ORIGIN           7 a 10 c 11 g 9 t

Query Match      0.8%; Score 18; DB 29; Length 37;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCGCTGCAGG 2342
Db 37 AGCTTGCATGCGCTGCAGG 20

RESULT 15
LOCUS            DR35H17T/c
DEFINITION      Danio rerio genomic clone DKEY-35H17, genomic survey sequence.
ACCESSION       AL975957
VERSION         AL975957.1  GI:25179324
KEYWORDS        GSS.
SOURCE          Danio rerio (zebrafish)
ORGANISM        Danio rerio
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
                Cypriniformes; Cyprinidae; Danio.
REFERENCE       1 (bases 1 to 43)
AUTHORS         Humphray,S.J., Huckle,E. and Hunt,S.E.
TITLE           Direct Submission
JOURNAL         Submitted (14-NOV-2002) The Sanger Institute, Wellcome Trust Genome
                Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
                humquerry@sanger.ac.uk Unpublished
COMMENT         This sequence was generated from the T7 end of BAC 35H17. 35H17 is
                part of the Daniokey BAC Library created by R. Plasterk and N.V.
                Keygene.
                Further details: http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES        Location/Qualifiers

```

source

```
1. .43
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-35H17"
/tissue_type="Testis"
/note="Vector pIndigoBAC-536"
7 a 11 c 12 g 13 t
```

BASE COUNT
ORIGIN

Query Match 0.8%; Score 18; DB 29; Length 43;
*Best Local Similarity 100.0%; Pred. No. 2.6e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCCTGCAGG 2342
|||
Db 37 AGCTTGCATGCCTGCAGG 20

Search completed: November 14, 2003, 19:05:45
Job time : 4906.81 secs


```
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/721,458B
; FILING DATE:
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/404,731
; FILING DATE: 15-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,227
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/503,236
; FILING DATE: 17-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/33276
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-4740448
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-721-458B-45

Query Match          0.9%; Score 21; DB 3; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.9; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGCG 2345
Db 29 AGCTTGCATGCTGCAGGTGCG 9

RESULT 3
US-09-039-198A-11/c
; Sequence 11, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-09-039-198A-11

Query Match          0.9%; Score 21; DB 3; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.9; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGCG 2345
Db 29 AGCTTGCATGCTGCAGGTGCG 9

RESULT 4
US-09-423-742-1/c
; Sequence 1, Application US/09423742
; Patent No. 6313284
; GENERAL INFORMATION:
; APPLICANT: Kwiatkowski, Marek
; APPLICANT: Landegren, Ulf
; APPLICANT: Nilsson, Mats
; TITLE OF INVENTION: SOLID PHASE SYNTHESIS
; FILE REFERENCE: 11989-005001
; CURRENT APPLICATION NUMBER: US/09/423,742
; CURRENT FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: PCT/SE98/00893
; PRIOR FILING DATE: 1998-05-14
; PRIOR APPLICATION NUMBER: SE 9701783-4
; PRIOR FILING DATE: 1997-05-14
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic oligonucleotide template of M13-30
; OTHER INFORMATION: complementary to the synthesized sequence
; US-09-423-742-1

Query Match          0.9%; Score 21; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.9; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGCG 2345
Db 23 AGCTTGCATGCTGCAGGTGCG 3

RESULT 5
US-08-877-599-11/c
; Sequence 11, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

; APPLICATION NUMBER: US/08/877,599
; FILING DATE: 514
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,618
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/33994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-877-599-11

Query Match 0.9%; Score 21; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTCGAGGTCG 2345
|||||
Db 29 AGCTTGCATGCTCGAGGTCG 9

RESULT 6

US-09-568-527-15
; Sequence 15, Application US/09568527
; Patent No. 6385080
; GENERAL INFORMATION:
; APPLICANT: HELLER, MICHAEL J.
; TITLE OF INVENTION: DNA OPTICAL STORAGE
; FILE REFERENCE: DAVID B. MURPHY - Heller 253/078
; CURRENT APPLICATION NUMBER: US/09/568,527
; CURRENT FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: 09/129,740
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 30
; TYPE: DNA
; ORGANISM: synthetic construct
US-09-568-527-15

Query Match 0.9%; Score 21; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTCGAGGTCG 2345
|||||
Db 5 AGCTTGCATGCTCGAGGTCG 25

RESULT 7

US-09-267-574-11/c
; Sequence 11, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/09/267,574
; CURRENT FILING DATE: 1999-03-12

; EARLIER APPLICATION NUMBER: 09/039,198
; EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer
US-09-267-574-11

Query Match 0.9%; Score 21; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTCGAGGTCG 2345
|||||
Db 29 AGCTTGCATGCTCGAGGTCG 9

RESULT 8

US-09-081-576-2/c
; Sequence 2, Application US/09081576A
; Patent No. 6180111
; GENERAL INFORMATION:
; APPLICANT: STEIN, Daniel C.
; TITLE OF INVENTION: VACCINE DELIVERY SYSTEM
; FILE REFERENCE: 2747-0097-27 CIP (8014-014)
; CURRENT APPLICATION NUMBER: US/09/081,576A
; CURRENT FILING DATE: 1998-05-19
; PRIOR APPLICATION NUMBER: US 08/936,522
; PRIOR FILING DATE: 1997-09-23
; PRIOR APPLICATION NUMBER: US 08/443,514
; PRIOR FILING DATE: 1995-05-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 32
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-09-081-576-2

Query Match 0.9%; Score 21; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTCGAGGTCG 2345
|||||
Db 29 AGCTTGCATGCTCGAGGTCG 9

RESULT 9

PCT-US95-14639-3/c
; Sequence 3, Application PC/TUS9514639
; GENERAL INFORMATION:
; APPLICANT: PHILLIPA MARRACK
; APPLICANT: JOHN KAPPLER
; APPLICANT: RICHARD SHIMONKEVITZ
; APPLICANT: MASAZUMI MATSUMURA
; TITLE OF INVENTION: PROTECTIVE EFFECTS OF MUTATED
; TITLE OF INVENTION: SUPERANTIGENS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Ave., Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; COMPUTER: IBM pc compatible

OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6.0 for windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/14639
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/338,373
FILING DATE: 14-November-1994
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: SUP020/PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US95-14639-3

Query Match 0.9%; Score 21; DB 5; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCCTGCAGGTGC 2345
DB 32 AGCTTGCATGCCTGCAGGTGC 12

RESULT 10
US-09-029-045-4
Sequence 4, Application US/09029045
Patent No. 6056952
GENERAL INFORMATION:
APPLICANT: Rosenberg, Amy Sonya
TITLE OF INVENTION: Selective Elimination of T Cells That
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/029,045
FILING DATE: 02-JUN-1998
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/002,964
FILING DATE: 30-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US96/13873
FILING DATE: 29-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 015280-236100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-029-045-4

Query Match 0.9%; Score 21; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCCTGCAGGTGC 2345
DB 14 AGCTTGCATGCCTGCAGGTGC 34

RESULT 11
US-09-435-321-4
Sequence 4, Application US/09435321-
Patent No. 6491908
GENERAL INFORMATION:
APPLICANT: Rosenberg, Amy Sonya
TITLE OF INVENTION: Selective Elimination of T Cells That
Recognize Specific Presselected Targets
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/435,321
FILING DATE: 04-No. 6491908-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/029,045
FILING DATE: 02-JUN-1998
APPLICATION NUMBER: US 60/002,964
FILING DATE: 30-AUG-1995
APPLICATION NUMBER: WO PCT/US96/13873
FILING DATE: 29-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 015280-236100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-435-321-4

Query Match 0.9%; Score 21; DB 4; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCCTGCAGGTGC 2345
DB 14 AGCTTGCATGCCTGCAGGTGC 34

RESULT 12
 US-09-387-300-39
 ; Sequence 39, Application US/09387300
 ; Patent No. 6358685
 ; GENERAL INFORMATION:
 ; APPLICANT: Wetmur, James G
 ; APPLICANT: Quatrin, Robin S
 ; APPLICANT: Engelhardt, Dean L
 ; TITLE OF INVENTION: Branch Migration of Nucleotides
 ; FILE REFERENCE: ENZ-49 (P) (C) SEQUENCES
 ; CURRENT APPLICATION NUMBER: US/09/387.300
 ; CURRENT FILING DATE: 1999-08-31
 ; EARLIER APPLICATION NUMBER: 08/480,000
 ; EARLIER FILING DATE: 1995-06-07
 ; NUMBER OF SEQ ID NOS: 39
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 39
 ; LENGTH: 37
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: MedC modified
 ; OTHER INFORMATION: complement to plasmid pMS19
 US-09-387-300-39

Query Match 0.9%; Score 21; DB 4; Length 37;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2325 AGCTTGCATGCTGCAGGTGC 2345
 Db 3 AGCTTGCATGCTGCAGGTGC 23

RESULT 13
 US-09-000-062-11/c
 ; Sequence 11, Application US/09000062
 ; Patent No. 6338961
 ; GENERAL INFORMATION:
 ; APPLICANT: DEROSE, Richard
 ; APPLICANT: CHAUBET, Nicole
 ; APPLICANT: GIGOT, Claude (deceased)
 ; TITLE OF INVENTION: ISOLATED DNA SEQUENCE CAPABLE OF SERVING AS REGULATORY
 ; TITLE OF INVENTION: ELEMENT IN A CHIMERIC GENE WHICH CAN BE USED FOR THE
 ; FILE REFERENCE: 022650-453
 ; CURRENT APPLICATION NUMBER: US/09/000.062
 ; CURRENT FILING DATE: 1998-05-29
 ; EARLIER APPLICATION NUMBER: PCT/FR96/01109
 ; EARLIER FILING DATE: 1996-07-17
 ; EARLIER APPLICATION NUMBER: FR 95/08980
 ; EARLIER FILING DATE: 1995-07-19
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 11
 ; LENGTH: 38
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 US-09-000-062-11

Query Match 0.9%; Score 21; DB 4; Length 38;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2325 AGCTTGCATGCTGCAGGTGC 2345
 Db 37 AGCTTGCATGCTGCAGGTGC 17

RESULT 14
 US-08-945-144A-8/c
 ; Sequence 8, Application US/08945144A

Patent No. 6566587
 ; GENERAL INFORMATION:
 ; APPLICANT: Lebrun, Michel
 ; APPLICANT: Salland, Alain
 ; APPLICANT: Freyssinet, Georges
 ; APPLICANT: Degryse, Eric
 ; TITLE OF INVENTION: Mutated 5-Enol Pyruvylshikimate-3-Phosphate Synthase,
 ; TITLE OF INVENTION: Gene Coding for Said Protein and Transformed Plants
 ; FILE REFERENCE: 5500-13
 ; CURRENT APPLICATION NUMBER: US/08/945.144A
 ; CURRENT FILING DATE: 1998-01-20
 ; PRIOR APPLICATION NUMBER: PCT/FR96/01125
 ; PRIOR FILING DATE: 1996-07-18
 ; PRIOR APPLICATION NUMBER: FRANCE 95/08979
 ; PRIOR FILING DATE: 1995-07-19
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 8
 ; LENGTH: 38
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 US-08-945-144A-8

Query Match 0.9%; Score 21; DB 4; Length 38;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2325 AGCTTGCATGCTGCAGGTGC 2345
 Db 37 AGCTTGCATGCTGCAGGTGC 17

RESULT 15
 US-08-425-684-78
 ; Sequence 78, Application US/08425684
 ; Patent No. 5834252
 ; GENERAL INFORMATION:
 ; APPLICANT: STEMMER PH.D., WILLEM P.C.
 ; TITLE OF INVENTION: END-COMPLEMENTARY POLYMERASE REACTION
 ; NUMBER OF SEQUENCES: 136
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW
 ; STREET: ONE MARKET PLAZA, STEUART STREET TOWER, 20TH FLOOR
 ; CITY: SAN FRANCISCO
 ; STATE: CALIFORNIA
 ; COUNTRY: U.S.A.
 ; ZIP: 94105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/425.684
 ; FILING DATE: 18-APR-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: DUNN ESQ., TRACY J.
 ; REGISTRATION NUMBER: 34,587
 ; REFERENCE/DOCKET NUMBER: 16528J-015400US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 326-2400
 ; TELEFAX: (415) 326-2422 78;
 ; INFORMATION FOR SEQ ID NO: 78:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 40 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: circular
 US-08-425-684-78

Query Match 0.9%; Score 21; DB 2; Length 40;

Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy . 2325 AGCTTGCATGCCCTGCAGGTTCG 2345
 |||||
Db 15 AGCTTGCATGCCCTGCAGGTTCG 35

Search completed: November 14, 2003, 16:31:52
Job time : 136.074 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 11:39:49 ; Search time 711.869 Seconds
(without alignments)
10825.622 Million cell updates/sec

Title: US-10-005-337A-1

Perfect score: 2358

Sequence: 1 ggaatcttctatgtttaaca.....caggtcgaggccaccatgg 2358

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 2169961 seqs, 1634102185 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1195920

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
- 12: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	0.9	27	13	US-10-086-464-26
2	21	0.9	30	9	US-09-952-944-1
3	21	0.9	30	12	US-10-161-547-11
4	21	0.9	32	13	US-10-153-902-2
5	21	0.9	38	14	US-10-023-839-11
6	21	0.9	40	9	US-09-245-802-78
7	21	0.9	46	11	US-09-940-925A-161
8	21	0.9	46	11	US-09-941-193A-161
9	21	0.9	50	13	US-10-106-832-2
10	20	0.8	24	9	US-09-231-235-10
11	20	0.8	24	9	US-09-797-518A-10
12	20	0.8	24	11	US-09-996-073-8
13	20	0.8	24	11	US-09-872-696A-10
14	20	0.8	28	11	US-09-123-064-5
15	19	0.8	20	11	US-09-949-427-252
16	19	0.8	25	9	US-09-879-257A-12

17	18	0.8	30	10	US-09-993-502-4	Sequence 4, Appl
18	18	0.8	32	9	US-09-231-235-25	Sequence 25, Appl
19	18	0.8	32	9	US-09-797-518A-25	Sequence 25, Appl
20	18	0.8	32	11	US-09-872-696A-25	Sequence 25, Appl
21	18	0.8	36	14	US-10-131-591A-56	Sequence 56, Appl
22	17	0.7	23	14	US-10-242-822B-27	Sequence 27, Appl
23	17	0.7	30	11	US-09-741-744A-100	Sequence 100, Appl
24	17	0.7	34	12	US-10-128-587A-11	Sequence 11, Appl
25	17	0.7	34	14	US-10-128-590-11	Sequence 11, Appl
26	17	0.7	50	11	US-09-943-007A-7	Sequence 7, Appl
27	16	0.7	17	13	US-10-021-237-2	Sequence 2, Appl
28	16	0.7	18	11	US-09-949-427-109	Sequence 109, Appl
29	16	0.7	26	11	US-09-776-191-30	Sequence 30, Appl
30	16	0.7	27	14	US-10-085-908-78	Sequence 78, Appl
31	16	0.7	29	12	US-10-302-983-6	Sequence 6, Appl
32	16	0.7	29	12	US-10-302-997-6	Sequence 6, Appl
33	16	0.7	30	11	US-09-940-244-20	Sequence 20, Appl
34	16	0.7	30	11	US-09-940-925A-20	Sequence 20, Appl
35	16	0.7	30	11	US-09-982-667-20	Sequence 20, Appl
36	16	0.7	30	11	US-09-941-193A-20	Sequence 20, Appl
37	16	0.7	30	12	US-10-290-386-20	Sequence 20, Appl
38	16	0.7	30	13	US-10-033-297-20	Sequence 20, Appl
39	16	0.7	30	13	US-10-081-806-20	Sequence 20, Appl
40	16	0.7	30	14	US-10-074-328-20	Sequence 20, Appl
41	16	0.7	31	10	US-09-995-593A-38	Sequence 38, Appl
42	16	0.7	36	11	US-09-776-191-43	Sequence 43, Appl
43	16	0.7	40	11	US-09-988-462-82	Sequence 82, Appl
44	16	0.7	40	12	US-10-170-172-41	Sequence 41, Appl
45	16	0.7	42	11	US-09-940-244-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-10-086-464-26
; Sequence 26, Application US/10086464
; Publication No. US20020199218A1
; GENERAL INFORMATION:
; APPLICANT: GORING, Daphne R. et al.
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
; FILE REFERENCE: P 25,762-A USA
; CURRENT APPLICATION NUMBER: US/10/086,464
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 10/069,304
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: PCT/CA00/00966
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/149,466
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: US 60/159,122
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 27
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer
US-10-086-464-26

Query Match 0.9%; Score 21; DB 13; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGCG 2345

Db 5 AGCTTGCATGCTGCAGGTGCG 25

RESULT 2

US-09-952-944-1/c

```
; Sequence 1, Application US/09952944
; Patent No. US20020051994A1
; GENERAL INFORMATION:
; APPLICANT: Kwiatkowski, Marek
; APPLICANT: Landegren, Ulf
; APPLICANT: Nilesen, Mats
; TITLE OF INVENTION: SOLID PHASE SYNTHESIS
; FILE REFERENCE: 11989-005001
; CURRENT APPLICATION NUMBER: US/09/952,944
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US/09/423,742
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: SE 9701783-4
; PRIOR FILING DATE: 1997-05-14
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic oligonucleotide template of M13-30
; OTHER INFORMATION: complementary to the synthesized sequence
US-09-952-944-1

Query Match          0.9%; Score 21; DB 9; Length 30;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTCG 2345
Db 23 AGCTTGCATGCTGCAGGTCG 3

RESULT 3
US-10-161-547-11/c
; Sequence 11, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/10/161,547
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/267,574
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer
US-10-161-547-11

Query Match          0.9%; Score 21; DB 12; Length 30;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTCG 2345
Db 29 AGCTTGCATGCTGCAGGTCG 9

RESULT 4
US-10-153-902-2/c
; Sequence 2, Application US/10153902
; Publication No. US20020187160A1
; GENERAL INFORMATION:
; APPLICANT: STEIN, Daniel C.
; TITLE OF INVENTION: VACCINE DELIVERY SYSTEM
; FILE REFERENCE: 2747-0097-27 CIP (8014-014)
; CURRENT APPLICATION NUMBER: US/10/153,902
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US/09/081,576
; PRIOR FILING DATE: 1998-05-19
; PRIOR APPLICATION NUMBER: US 08/936,522
; PRIOR FILING DATE: 1997-09-23
; PRIOR APPLICATION NUMBER: US 08/443,514
; PRIOR FILING DATE: 1995-05-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 32
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-153-902-2

Query Match          0.9%; Score 21; DB 13; Length 32;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTCG 2345
Db 29 AGCTTGCATGCTGCAGGTCG 9

RESULT 5
US-10-023-839-11/c
; Sequence 11, Application US/10023839
; Publication No. US20030027312A1
; GENERAL INFORMATION:
; APPLICANT: DEROSE, Richard
; APPLICANT: CHAUBET, Nicole
; TITLE OF INVENTION: ISOLATED DNA SEQUENCE CAPABLE OF SERVING AS REGULATORY
; TITLE OF INVENTION: ELEMENT IN A CHIMERIC GENE WHICH CAN BE USED FOR THE
; TITLE OF INVENTION: TRANSFORMATION OF PLANTS
; FILE REFERENCE: 022650-453
; CURRENT APPLICATION NUMBER: US/10/023,839
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/000,062
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: FR 95/08980
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-07-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 38
; TYPE: DNA
; ORGANISM: Zea mays
US-10-023-839-11

Query Match          0.9%; Score 21; DB 14; Length 38;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTCG 2345
Db 37 AGCTTGCATGCTGCAGGTCG 17

RESULT 6
US-09-245-802-78
; Sequence 78, Application US/09245802
; Patent No. US20010049125A1
; GENERAL INFORMATION:
; APPLICANT: STEMMER PH.D., WILLEM P.C.
; APPLICANT: LIPSCHUTZ, ROBERT J.
; TITLE OF INVENTION: END-COMPLEMENTARY POLYMERASE REACTION
; NUMBER OF SEQUENCES: 136
; CORRESPONDENCE ADDRESS:
```



```
;
; ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW
; STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: U.S.A.
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/245,802
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,502
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05480
; FILING DATE: 18-APR-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY, ESQ., MATTHEW B.
; REGISTRATION NUMBER: 39,787
; REFERENCE/DOCKET NUMBER: 16528J-015410US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 78:
; LENGTH: 40 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; US-09-245-802-78

Query Match 0.9%; Score 21; DB 9; Length 40;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTGCATGCTGCAGGTGC 2345
Db 15 AGCTGCATGCTGCAGGTGC 35

RESULT 7
US-09-940-925A-161
; Sequence 161, Application US/09940925A
; Publication No. US20030054338A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/940,925A
; FILING DATE: 10-JUN-2002

Query Match 0.9%; Score 21; DB 9; Length 40;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTGCATGCTGCAGGTGC 2345
Db 15 AGCTGCATGCTGCAGGTGC 35

RESULT 7
US-09-940-925A-161
; Sequence 161, Application US/09940925A
; Publication No. US20030054338A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/940,925A
; FILING DATE: 10-JUN-2002
```

```
;
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 161:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 161:
US-09-940-925A-161

Query Match 0.9%; Score 21; DB 11; Length 46;
Best Local Similarity 76.2%; Pred. No. 6.6;
Matches 16; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTGCATGCTGCAGGTGC 2345
Db 10 AGCUUGCAUGCCUGCAGGUCG 30

RESULT 8
US-09-941-193A-161
; Sequence 161, Application US/09941193A
; Publication No. US20030108873A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; LYAMICHEV, VICTOR I.
; OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/941,193A
; FILING DATE: 28-Aug-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 161:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 161:
US-09-941-193A-161

Query Match 0.9%; Score 21; DB 11; Length 46;
Best Local Similarity 76.2%; Pred. No. 6.6;
```

Matches 16; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGC 2345
Db 10 AGCUUGCAUGCCGCGAGGUCG 30

RESULT 9
US-10-106-832-2/c
; Sequence 2, Application US/10106832
; Publication No. US20020177698A1
; GENERAL INFORMATION:
; APPLICANT: Kwiatkowski, Marek
; APPLICANT: Landegren, Ulf
; APPLICANT: Nilsson, Mats
; TITLE OF INVENTION: SOLID PHASE SYNTHESIS
; FILE REFERENCE: 11989-006001
; CURRENT APPLICATION NUMBER: US/10/106,832
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 09/254,048
; PRIOR FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: circularization probe
US-10-106-832-2

Query Match 0.9%; Score 21; DB 13; Length 50;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTGC 2345
Db 33 AGCTTGCATGCTGCAGGTGC 13

RESULT 10
US-09-231-235-10
; Sequence 10, Application US/09231235
; Patent No. US20020048805A1
; GENERAL INFORMATION:
; APPLICANT: Johnston, Julie C.
; APPLICANT: Sauter, Sybille L.
; APPLICANT: Hsu, David
; APPLICANT: Sheridan, Philip Lee
; APPLICANT: Hardy, Steven
; APPLICANT: Dubensky, Thomas
; APPLICANT: Yee, Jiling-Kuan
; TITLE OF INVENTION: FELINE IMMUNODEFICIENCY VIRUS GENE THERAPY VECTORS
; FILE REFERENCE: 930049.467
; CURRENT APPLICATION NUMBER: US/09/231,235
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
; OTHER INFORMATION: Primer Used in Examples 2A and 2B
US-09-231-235-10

Query Match 0.8%; Score 20; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2322 GGGAGCTTGCATGCTGCAG 2341

Db 5 GGGAGCTTGCATGCTGCAG 24

RESULT 11
US-09-797-518A-10
; Sequence 10, Application US/09797518A
; Patent No. US20020068354A1
; GENERAL INFORMATION:
; APPLICANT: Johnston, Julie C.
; APPLICANT: Sauter, Sybille L.
; APPLICANT: Hsu, David
; APPLICANT: Sheridan, Philip Lee
; APPLICANT: Hardy, Steven
; APPLICANT: Dubensky, Thomas
; APPLICANT: Yee, Jiling-Kuan
; TITLE OF INVENTION: FELINE IMMUNODEFICIENCY VIRUS GENE THERAPY VECTORS
; FILE REFERENCE: 930049.467
; CURRENT APPLICATION NUMBER: US/09/797,518A
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/231,235
; PRIOR FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
; OTHER INFORMATION: Primer Used in Examples 2A and 2B
US-09-797-518A-10

Query Match 0.8%; Score 20; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2322 GGGAGCTTGCATGCTGCAG 2341
Db 5 GGGAGCTTGCATGCTGCAG 24

RESULT 12
US-09-996-073-8
; Sequence 8, Application US/09996073
; Publication No. US20030003565A1
; GENERAL INFORMATION:
; APPLICANT: DUBENSKY JR., Thomas W.
; APPLICANT: GASMI, Mehdi
; APPLICANT: SAUTER, Sybille
; TITLE OF INVENTION: FUNCTIONAL LENTIVIRAL VECTOR FROM
; FILE REFERENCE: 2302-1642 / 1642.002
; CURRENT APPLICATION NUMBER: US/09/996,073
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: FIV19 primer
US-09-996-073-8

Query Match 0.8%; Score 20; DB 11; Length 24;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2322 GGGAGCTTGCATGCTGCAG 2341
Db 5 GGGAGCTTGCATGCTGCAG 24

RESULT 13

US-09-872-696A-10

; Sequence 10, Application US/09872696A

; Publication No. US20030104611A1

; GENERAL INFORMATION:

; APPLICANT: Johnston, Julie C.

; APPLICANT: Sauter, Sybille L.

; APPLICANT: Heu, David

; APPLICANT: Sheridan, Philip Lee

; APPLICANT: Hardy, Steven

; APPLICANT: Dubensky, Thomas

; APPLICANT: Yee, Jiling-Kuan

; TITLE OF INVENTION: FELINE IMMUNODEFICIENCY VIRUS GENE THERAPY VECTORS

; FILE REFERENCE: 930049.467

; CURRENT APPLICATION NUMBER: US/09/872.696A

; CURRENT FILING DATE: 1999-01-18

; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 10

; LENGTH: 24

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Primer

; OTHER INFORMATION: Primer Used in Examples 2A and 2B

US-09-872-696A-10

Query Match 0.8%; Score 20; DB 11; Length 24;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2322 GGGAGCTTGCATGCTGCAG 2341

Db 5 GGGAGCTTGCATGCTGCAG 24

RESULT 14

US-09-123-064-5

; Sequence 5, Application US/09123064

; Publication No. US20030082149A1

; GENERAL INFORMATION:

; APPLICANT: Rowe, David W.

; APPLICANT: Stover, Mary Louise

; APPLICANT: Beckley, Akin

; TITLE OF INVENTION: INHIBITION OF A TARGET MESSENGER RNA

; FILE REFERENCE: UCT97-02pA

; CURRENT APPLICATION NUMBER: US/09/123.064

; CURRENT FILING DATE: 1998-07-27

; EARLIER APPLICATION NUMBER: US 60/053,998

; EARLIER FILING DATE: 1997-07-28

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 5

; LENGTH: 28

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Oligonucleotide

US-09-123-064-5

Query Match 0.8%; Score 20; DB 11; Length 28;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 AGCTTGCATGCTGCAGGTC 2344

Db 9 AGCTTGCATGCTGCAGGTC 28

RESULT 15

US-09-949-427-252

; Sequence 252, Application US/09949427
; Publication No. US20030054418A1
; GENERAL INFORMATION:
; APPLICANT: Bodnar, Jackie S.
; APPLICANT: Castellani, Lawrence W.
; APPLICANT: Chatterjee, Aurobindo
; APPLICANT: de Jong, Pieter
; APPLICANT: Lusia, Aldons J.
; APPLICANT: Ohmen, Jeff
; APPLICANT: Ross, David
; APPLICANT: Tafuri, Sherrie
; APPLICANT: Wu, Chenyan
; TITLE OF INVENTION: Gene and Sequence Variation Associated with Cancer
; FILE REFERENCE: 02810.0014.NPUS02
; CURRENT APPLICATION NUMBER: US/09/949,427
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/231,322
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 252
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Primer
US-09-949-427-252

Query Match 0.8%; Score 19; DB 11; Length 20;

Best Local Similarity 100.0%; Pred. No. 72;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2327 CTTGCATGCTGCAGGTCG 2345

Db 1 CTTGCATGCTGCAGGTCG 19

Search completed: November 14, 2003, 19:28:14

Job time : 715.869 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 14, 2003, 06:44:53 ; Search time 7536.5 Seconds
(without alignments)
11258.085 Million cell updates/sec

Title: US-10-005-337A-2
Perfect score: 2074
Sequence: 1 ctgcagcaagttacttaatg.....acaagactcttcagccaac 2074

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1316618

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rnd.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	18	0.9	34	6	I39756	I39756 Sequence 29
C 2	17	0.8	24	6	AX084499	AX084499 Sequence
C 3	16	0.8	18	6	AR106883	AR106883 Sequence
C 4	16	0.8	19	6	AR293216	AR293216 Sequence
C 5	16	0.8	43	6	AX484404	AX484404 Sequence
C 6	16	0.8	47	6	AR290608	AR290608 Sequence
C 7	16	0.8	47	6	AR290806	AR290806 Sequence
C 8	16	0.8	48	9	HSU08249	U08249 Human chrom
C 9	16	0.8	48	10	MMHOX31B	X12653 Mouse Hox-3
C 10	15	0.7	15	6	AR041401	AR041401 Sequence
C 11	15	0.7	15	6	AR041402	AR041402 Sequence
C 12	15	0.7	15	6	AR041409	AR041409 Sequence
C 13	15	0.7	15	6	AR041410	AR041410 Sequence
C 14	15	0.7	15	6	AR041411	AR041411 Sequence
C 15	15	0.7	15	6	AR041922	AR041922 Sequence
C 16	15	0.7	15	6	AR041923	AR041923 Sequence
C 17	15	0.7	15	6	AR041924	AR041924 Sequence
C 18	15	0.7	15	6	AR041925	AR041925 Sequence
C 19	15	0.7	15	6	AR041926	AR041926 Sequence
C 20	15	0.7	15	6	AR041931	AR041931 Sequence
C 21	15	0.7	15	6	AX636860	AX636860 Sequence
C 22	15	0.7	15	6	AX636862	AX636862 Sequence
C 23	15	0.7	15	6	AX636876	AX636876 Sequence
C 24	15	0.7	15	6	AX636878	AX636878 Sequence
C 25	15	0.7	15	6	AX636880	AX636880 Sequence
C 26	15	0.7	15	6	AX637393	AX637393 Sequence
C 27	15	0.7	15	6	AX637395	AX637395 Sequence
C 28	15	0.7	15	6	AX637397	AX637397 Sequence
C 29	15	0.7	15	6	AX637399	AX637399 Sequence
C 30	15	0.7	15	6	AX637401	AX637401 Sequence
C 31	15	0.7	15	6	AX637411	AX637411 Sequence
C 32	15	0.7	17	6	AX672994	AX672994 Sequence
C 33	15	0.7	17	6	AX738592	AX738592 Sequence
C 34	15	0.7	20	6	AB3584	AB3584 Sequence 13
C 35	15	0.7	20	6	AB3598	AB3598 Sequence 27
C 36	15	0.7	20	6	AR004674	AR004674 Sequence
C 37	15	0.7	20	6	AR008160	AR008160 Sequence
C 38	15	0.7	20	6	AR136943	AR136943 Sequence
C 39	15	0.7	20	6	AR224513	AR224513 Sequence
C 40	15	0.7	20	6	AR228975	AR228975 Sequence
C 41	15	0.7	20	6	AR242931	AR242931 Sequence
C 42	15	0.7	20	6	AX384983	AX384983 Sequence
C 43	15	0.7	20	6	BD105584	BD105584 Genes sen
C 44	15	0.7	20	6	I76944	I76944 Sequence 3
C 45	15	0.7	20	6	I80939	I80939 Sequence 3

ALIGNMENTS

RESULT 1
I39756/c
LOCUS I39756
DEFINITION Sequence 29 from patent US 5616490.
ACCESSION I39756
VERSION I39756.1 GI:2084236
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 34)
AUTHORS Sullivan,S.M. and Draper,K.G.
TITLE Ribozymes targeted to TNF- α . RNA
JOURNAL Patent: US 5616490-A 29 01-APR-1997;
FEATURES Location/Qualifiers

linear PAT 13-MAY-1997

```

source
1. .34
/organism="unknown"
BASE COUNT      10 a      0 c      0 g      24 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 34;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 AATAAATAAATAATAAA 397
|||||
Db 33 AATAAATAAATAATAAA 16

RESULT 2
LOCUS AX084499/c 24 bp mRNA linear PAT 28-FEB-2001
DEFINITION Sequence 41 from Patent WO0112213.
ACCESSION AX084499
VERSION AX084499.1 GI:13185910
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
TITLE Blackshear,P.J., Lai,W.S. and Carballo-Jane,E.
JOURNAL Ttp-related zinc finger domains and methods of use
FEATURES Patent: WO 0112213-A 41 22-FEB-2001;
THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
Location/Qualifiers
source
1. .24
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
18 t
BASE COUNT      6 a      0 c      0 g      18 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 17; DB 6; Length 24;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 AATAAATAAATAATAAA 397
|||||
Db 24 AATAAATAAATAATAAA 8

RESULT 3
LOCUS AR106883 18 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 44 from patent US 6107092.
ACCESSION AR106883
VERSION AR106883.1 GI:12821413
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 18)
AUTHORS Cowsett,L.M., Bennett,C.Frank. and O'Malley,B.W.
TITLE Antisense modulation of SRA expression
JOURNAL Patent: US 6107092-A 44 22-AUG-2000;
FEATURES Location/Qualifiers
source
1. .18
/organism="unknown"
BASE COUNT      5 a      8 c      2 g      3 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 16; DB 6; Length 18;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 601 ACCATCCACTGACTGA 616
|||||

Db 1 ACCATCCACTGACTGA 16

RESULT 4
LOCUS AR293216 19 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 4951 from patent US 6537751.
ACCESSION AR293216
VERSION AR293216.1 GI:31680500
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 19)
AUTHORS Cohen,D., Chumakov,I. and Blumenfeld,M.
TITLE Biallelic markers for use in constructing a high density
disequilibrium map of the human genome
JOURNAL Patent: US 6537751-A 4951 25-MAR-2003;
FEATURES Location/Qualifiers
source
1. .19
/organism="unknown"
BASE COUNT      10 a      5 c      2 g      2 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 16; DB 6; Length 19;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2046 GAAAAACATACAGAC 2061
|||||
Db 2 GAAAAACATACAGAC 17

RESULT 5
LOCUS AX484404 43 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 1704 from Patent WO02053728.
ACCESSION AX484404
VERSION AX484404.1 GI:22318756
KEYWORDS Candida albicans
SOURCE Candida albicans
ORGANISM Candida albicans
REFERENCE 1
AUTHORS Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
TITLE Gene disruption methodologies for drug target discovery
JOURNAL Patent: WO 02053728-A 1704 11-JUL-2002;
Elitra Pharmaceuticals, Inc. (US)
Location/Qualifiers
source
1. .43
/organism="Candida albicans"
/mol_type="genomic DNA"
/db_xref="taxon:5476"
25 t
BASE COUNT      12 a      5 c      1 g      25 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 16; DB 6; Length 43;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 AAAAAATAAATAATAAA 393
|||||
Db 22 AAAAAATAAATAATAAA 7

RESULT 6
LOCUS AR290608 47 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 2343 from patent US 6537751.
ACCESSION AR290608
VERSION AR290608.1 GI:31677892
KEYWORDS

```

```

SOURCE      Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 47)
AUTHORS      Cohen,D., Chumakov,I. and Blumenfeld,M.
TITLE        Biallelic markers for use in constructing a high density
              disequilibrium map of the human genome
JOURNAL      Patent: US 6537751-A 2343 25-MAR-2003;
FEATURES     Location/Qualifiers
              1..47
              /organism="unknown"
BASE COUNT   33 a 3 c 1 g 9 t 1 others
ORIGIN
Query Match 0.8%; Score 16; DB 6; Length 47;
Best Local Similarity 100.0%; Pred. No. 2.1e+04;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 AAAAATAAATAATAA 393
Db 6 AAAAATAAATAATAA 21

RESULT 7
AR290806/c
LOCUS        AR290806 47 bp DNA linear PAT 12-JUN-2003
DEFINITION   Sequence 2541 from patent US 6537751.
ACCESSION    AR290806
VERSION      AR290806.1 GI:31678090
KEYWORDS     Unknown.
SOURCE       Unclassified.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 47)
AUTHORS      Cohen,D., Chumakov,I. and Blumenfeld,M.
TITLE        Biallelic markers for use in constructing a high density
              disequilibrium map of the human genome
JOURNAL      Patent: US 6537751-A 2541 25-MAR-2003;
FEATURES     Location/Qualifiers
              1..47
              /organism="unknown"
BASE COUNT   8 a 12 c 16 g 10 t 1 others
ORIGIN
Query Match 0.8%; Score 16; DB 6; Length 47;
Best Local Similarity 100.0%; Pred. No. 2.1e+04;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1253 AGCCACTGCACCCGGC 1268
Db 21 AGCCACTGCACCCGGC 6

RESULT 8
HSU08249/c
LOCUS        HSU08249 48 bp DNA linear PRI 29-APR-1994
DEFINITION   Human chromosome 1 clone pFU335R-S2.
ACCESSION    U08249
VERSION      U08249.1 GI:475536
KEYWORDS     Homo sapiens (human)
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 48)
AUTHORS      Brodeur,G.M.
TITLE        FUCAL YAC FU335 right insert end sequence #2 of 2
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 48)
AUTHORS      Brodeur,G.M.
TITLE        Direct Submission
JOURNAL      Submitted (31-MAR-1994) Garrett M. Brodeur, Oncology, Children's
              Hospital of Philadelphia, 34th and Civic Center Blvd, Philadelphia,

SOURCE      PA 19104, USA
FEATURES     Location/Qualifiers
              1..48
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="1"
              /map="1p36"
              /clone="pFU335R-S2"
              /sex="male"
              /clone_lib="CEPH YAC library, clone 335f4"
BASE COUNT   11 a 2 c 6 g 29 t
ORIGIN
Query Match 0.8%; Score 16; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.1e+04;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 AAAAATAAATAATAA 393
Db 38 AAAAATAAATAATAA 23

RESULT 9
MMH0X31B
LOCUS        MMH0X31B 48 bp mRNA linear ROD 08-JUL-1992
DEFINITION   Mouse Hox-3.1 homeo gene 3' flank.
ACCESSION    X12653 Y00215
VERSION      X12653.1 GI:51405
KEYWORDS     homeobox; Hox-3.1 gene.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
REFERENCE    1 (bases 1 to 48)
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
TITLE        Le Mouellic,H., Condamine,H. and Brulet,P.
              Pattern of transcription of the homeo gene Hox-3.1 in the mouse
              embryo
JOURNAL      Genes Dev. 2 (1), 125-135 (1988)
MEDLINE      88185818
PUBMED       2895723
COMMENT      see x12652 for Hox-3.1 gene 5' flank; see x07439 for Hox-3.1 cDNA
              sequence; the sequences are represented contiguously by [1].
FEATURES     Location/Qualifiers
              1..48
              /organism="Mus musculus"
              /mol_type="mRNA"
              /db_xref="taxon:10090"
              misc_feature 1..48
              /note="3' genomic flanking region"
BASE COUNT   32 a 1 c 11 g 4 t
ORIGIN
Query Match 0.8%; Score 16; DB 10; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.1e+04;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 408 AAAAATAAATAAGAAA 423
Db 29 AAAAATAAATAAGAAA 44

RESULT 10
AR041401/c
LOCUS        AR041401 15 bp DNA linear PAT 29-SEP-1999
DEFINITION   Sequence 191 from patent US 5811300.
ACCESSION    AR041401
VERSION      AR041401.1 GI:5961897
KEYWORDS     Unknown.
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 15)
AUTHORS      Sullivan,S., Draper,K., Kisich,K., Stinchcomb,D.T. and McSwiggen,J.

```

```

TITLE      TNF-.alpha. ribozymes
JOURNAL    Patent: US 5811300-A 191 22-SEP-1998;
FEATURES   Location/Qualifiers
           source
           1. .15
BASE COUNT 4 a 0 c 0 g 11 t
ORIGIN

Query Match      0.7%; Score 15; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.2e+04;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      382 ATAAATAAATAATAA 396
Db      15 ATAAATAAATAATAA 1

RESULT 11
AR041402/c
LOCUS      15 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 192 from patent US 5811300.
ACCESSION  AR041402
VERSION     AR041402.1 GI:5961898
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 15)
AUTHORS    Sullivan,S., Draper,K., Kisich,K., Stinchcomb,D.T. and McSwiggen,J.
TITLE      TNF-.alpha. ribozymes
JOURNAL    Patent: US 5811300-A 192 22-SEP-1998;
FEATURES   Location/Qualifiers
           source
           1. .15
BASE COUNT 4 a 0 c 0 g 11 t
ORIGIN

Query Match      0.7%; Score 15; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.2e+04;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      381 AATAAATAAATAATA 395
Db      15 AATAAATAAATAATA 1

RESULT 14
AR041411/c
LOCUS      15 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 201 from patent US 5811300.
ACCESSION  AR041411
VERSION     AR041411.1 GI:5961907
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 15)
AUTHORS    Sullivan,S., Draper,K., Kisich,K., Stinchcomb,D.T. and McSwiggen,J.
TITLE      TNF-.alpha. ribozymes
JOURNAL    Patent: US 5811300-A 201 22-SEP-1998;
FEATURES   Location/Qualifiers
           source
           1. .15
BASE COUNT 4 a 0 c 0 g 11 t
ORIGIN

Query Match      0.7%; Score 15; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.2e+04;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      380 AAATAAATAAATAAT 394
Db      15 AAATAAATAAATAAT 1

RESULT 15
AR041922/c
LOCUS      15 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 712 from patent US 5811300.
ACCESSION  AR041922
VERSION     AR041922.1 GI:5962418
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 15)
AUTHORS    Sullivan,S., Draper,K., Kisich,K., Stinchcomb,D.T. and McSwiggen,J.
TITLE      TNF-.alpha. ribozymes
JOURNAL    Patent: US 5811300-A 199 22-SEP-1998;
FEATURES   Location/Qualifiers
           source
           1. .15
BASE COUNT 4 a 0 c 0 g 11 t
ORIGIN

Query Match      0.7%; Score 15; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.2e+04;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      382 ATAAATAAATAATAA 396
Db      15 ATAAATAAATAATAA 1

```


JOURNAL Patent: US 5811300-A 712 22-SEP-1998;

FEATURES Location/Qualifiers

source 1. .15

/organism="unknown"

BASE COUNT 4 a 0 c 0 g 11 t

ORIGIN

Query Match 0.7%; Score 15; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.2e+04;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 382 ATAAATAATAATAA 396

Db 15 ATAAATAATAATAA 1

Search completed: November 14, 2003, 16:27:36
Job time : 7539.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 14, 2003, 06:44:53 ; Search time 8568.5 Seconds
(without alignments)
11258.085 Million cell updates/sec

Title: US-10-005-337A-1
Perfect score: 2358
Sequence: 1 ggaatcttctatgtttaaca.....caggtcgaggccaccatgg 2358

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 20454813386 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1316618

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hug.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_ats.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	0.9	48	6	A27635	A27635 JGMV coat p
2	21	0.9	22	6	AX012654	AX012654 Sequence
3	21	0.9	27	6	AX088898	AX088898 Sequence
4	21	0.9	30	6	AR107024	AR107024 Sequence
5	21	0.9	30	6	AR177937	AR177937 Sequence
6	21	0.9	30	6	AR206049	AR206049 Sequence
7	21	0.9	30	6	AR209748	AR209748 Sequence
8	21	0.9	30	6	AR212146	AR212146 Sequence
9	21	0.9	30	6	BD064156	BD064156 Chitinase
10	21	0.9	30	6	BD085067	BD085067 Synthesis
11	21	0.9	30	6	BD135123	BD135123 Chitin-Bi
12	21	0.9	30	6	BD178195	BD178195 Novel DNA
13	21	0.9	32	6	A23170	A23170 Artificial
14	21	0.9	32	6	AR126240	AR126240 Sequence
15	21	0.9	35	6	AR264638	AR264638 Sequence
16	21	0.9	37	6	AR200696	AR200696 Sequence
17	21	0.9	38	6	AR182676	AR182676 Sequence
18	21	0.9	38	6	AX463801	AX463801 Sequence
19	21	0.9	40	6	AR053654	AR053654 Sequence
20	21	0.9	40	6	AR070708	AR070708 Sequence
21	21	0.9	40	6	AR258589	AR258589 Sequence
22	21	0.9	40	6	E12434	E12434 Oligonucleo
23	21	0.9	48	6	A16034	A16034 oligonucleo
24	21	0.9	48	12	SYNECOMPB2	M24052 E.coli vect
25	21	0.9	50	6	AR035666	AR035666 Sequence
26	21	0.9	50	6	AR152918	AR152918 Sequence
27	21	0.9	50	6	AR170726	AR170726 Sequence
28	21	0.9	50	6	AR222506	AR222506 Sequence
29	21	0.9	50	6	E12581	E12581 Probe, 4/19
30	20	0.8	24	6	AX600313	AX600313 Sequence
31	20	0.8	43	6	E51368	E51368 Plasmid vec
32	20	0.8	45	6	A07929	A07929 Nucleotide
33	20	0.8	46	6	AR063274	AR063274 Sequence
34	20	0.8	46	12	S38358	S38358 {multiple c
35	20	0.8	49	6	A07930	A07930 Nucleotide
36	19	0.8	19	6	AX539200	AX539200 Sequence
37	19	0.8	20	6	AR020466	AR020466 Sequence
38	19	0.8	20	6	AX404665	AX404665 Sequence
39	19	0.8	20	6	AX477161	AX477161 Sequence
40	19	0.8	20	6	AX528537	AX528537 Sequence
41	19	0.8	22	6	A42926	A42926 Sequence 58
42	19	0.8	22	6	AR271460	AR271460 Sequence
43	19	0.8	22	6	AX404679	AX404679 Sequence
44	19	0.8	22	6	AX404682	AX404682 Sequence
45	19	0.8	25	6	AX383971	AX383971 Sequence

ALIGNMENTS

RESULT 1
A27635
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

A27635
JGMV coat protein N-terminal oligomer.
A27635
A27635.1 GI:1248477
synthetic construct
synthetic construct
artificial sequences.
1 (bases 1 to 48)
SELF-POLYMERISING EXPRESSION SYSTEM BASED ON MODIFIED POTYVIRUS
COAT PROTEINS
Patent: WO 9115587-A 7 17-OCT-1991;

48 bp
DNA
linear
PAT 25-SEP-1995

```
FEATURES
source
  Location/Qualifiers
    1..48
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"
BASE COUNT
  14 a 10 c 13 g 11 t
ORIGIN
Query Match
  0.9%; Score 22; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2324 GAGCTTCATGCGCTGCAGGTCG 2345
Db 3 GAGCTTCATGCGCTGCAGGTCG 24
RESULT 2
AX012654/c
LOCUS AX012654 22 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 49 from Patent WO9954490.
ACCESSION AX012654
VERSION AX012654.1 GI:9998632
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Baulcombe,D.C., Bendahmane,A. and Kanyuka,K.V.
TITLE Plant-derived resistance gene
JOURNAL Patent: WO 9954490-A 49 28-OCT-1999;
BAULCOMBE DAVID CHARLES (GB); BENDAHMANE ABDELHAFID (GB); KANYUKA
KONSTANTIN VALERIEVICH (GB); PLANT BIOSCIENCE LIMITED (GB)
FEATURES
source
  Location/Qualifiers
    1..22
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"
    /note="primer"
BASE COUNT
  5 a 7 c 6 g 4 t
ORIGIN
Query Match
  0.9%; Score 21; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2325 AGCTTGCATGCGCTGCAGGTCG 2345
Db 21 AGCTTGCATGCGCTGCAGGTCG 1
RESULT 3
AX088898
LOCUS AX088898 27 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 23 from Patent WO0114563.
ACCESSION AX088898
VERSION AX088898.1 GI:13397658
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Goring,D. and Silva,N.
TITLE Proline-rich extensin-like receptor kinases
JOURNAL Patent: WO 0114563-A 23 01-MAR-2001;
Goring, Daphne (CA); Silva, Nancy (CA)
FEATURES
source
  Location/Qualifiers
    1..27
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"
    /note="primer"
BASE COUNT
  6 a 7 c 9 g 5 t
ORIGIN
Query Match
  0.9%; Score 21; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2325 AGCTTGCATGCGCTGCAGGTCG 2345
Db 5 AGCTTGCATGCGCTGCAGGTCG 25
RESULT 4
AR107024/c
LOCUS AR107024 30 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 45 from patent US 6107104.
ACCESSION AR107024
VERSION AR107024.1 GI:12821554
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Lockerie,R.Owen., Howard,M.L., Gallatin,W.Michael. and Lai,Y.
TITLE Modulators of anchoring protein function
JOURNAL Patent: US 6107104-A 45 22-AUG-2000;
FEATURES
source
  Location/Qualifiers
    1..30
    /organism="unknown"
BASE COUNT
  7 a 8 c 8 g 7 t
ORIGIN
Query Match
  0.9%; Score 21; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2325 AGCTTGCATGCGCTGCAGGTCG 2345
Db 29 AGCTTGCATGCGCTGCAGGTCG 9
RESULT 5
AR177937/c
LOCUS AR177937 30 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 1 from patent US 6313284.
ACCESSION AR177937
VERSION AR177937.1 GI:17920292
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Kwiatkowski,M., Landegren,U. and Nilsson,M.
TITLE Solid phase synthesis
JOURNAL Patent: US 6313284-A 1 06-NOV-2001;
FEATURES
source
  Location/Qualifiers
    1..30
    /organism="unknown"
BASE COUNT
  6 a 9 c 9 g 6 t
ORIGIN
Query Match
  0.9%; Score 21; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2325 AGCTTGCATGCGCTGCAGGTCG 2345
Db 23 AGCTTGCATGCGCTGCAGGTCG 3
RESULT 6
AR206049/c
LOCUS AR206049 30 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 11 from patent US 6372212.
```

```

ACCESSION   AR206049
VERSION     AR206049.1  GI:21504535
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 30)
AUTHORS     Gray,P.W.
TITLE       Chitinase materials and methods
JOURNAL     Patent: US 6372212-A 11 16-APR-2002;
FEATURES    Location/Qualifiers
             1..30
             /organism="unknown"
BASE COUNT  7 a 8 c 8 g 7 t
ORIGIN

Query Match      0.9%; Score 21; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2325 AGCTTGCATGCTGCAGGTGC 2345
Db  29 AGCTTGCATGCTGCAGGTGC 9

RESULT 7
AR209748
LOCUS       AR209748             30 bp      DNA             linear      PAT 20-JUN-2002
DEFINITION Sequence 15 from patent US 6385080.
ACCESSION  AR209748
VERSION    AR209748.1  GI:21511254
KEYWORDS   Unknown.
SOURCE     Unknown.
REFERENCE   1 (bases 1 to 30)
AUTHORS     Heller,M.J. and Tu,E.
TITLE       DNA optical storage device using Forster energy transfer mechanism
JOURNAL     Patent: US 6385080-A 15 07-MAY-2002;
FEATURES    Location/Qualifiers
             1..30
             /organism="unknown"
BASE COUNT  5 a 10 c 8 g 7 t
ORIGIN

Query Match      0.9%; Score 21; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2325 AGCTTGCATGCTGCAGGTGC 2345
Db  5 AGCTTGCATGCTGCAGGTGC 25

RESULT 8
AR212146/c
LOCUS       AR212146             30 bp      DNA             linear      PAT 20-JUN-2002
DEFINITION Sequence 11 from patent US 6399571.
ACCESSION  AR212146
VERSION    AR212146.1  GI:21515653
KEYWORDS   Unknown.
SOURCE     Unknown.
REFERENCE   1 (bases 1 to 30)
AUTHORS     Gray,P.W. and Tjoelker,L.W.
TITLE       Chitinase chitin-binding fragments
JOURNAL     Patent: US 6399571-A 11 04-JUN-2002;
FEATURES    Location/Qualifiers
             1..30
             /organism="unknown"
BASE COUNT  7 a 8 c 8 g 7 t
ORIGIN

```

```

Query Match      0.9%; Score 21; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2325 AGCTTGCATGCTGCAGGTGC 2345
Db  29 AGCTTGCATGCTGCAGGTGC 9

RESULT 9
BD064156/c
LOCUS       BD064156             30 bp      DNA             linear      PAT 27-AUG-2002
DEFINITION Chitinase materials and methods.
ACCESSION  BD064156
VERSION    BD064156.1  GI:22609759
KEYWORDS   JP 2001510325-A/9.
SOURCE     synthetic construct
ORGANISM   synthetic construct
REFERENCE   1 (bases 1 to 30)
AUTHORS     Gray,P.W.
TITLE       Chitinase materials and methods
JOURNAL     Patent: JP 2001510325-A 9 31-JUL-2001;
COMMENT     ICOS CORP
             PN JP 2001510325-A/9
             PD 31-JUL-2001
             PF 16-JUN-1997 JP 1998501889
             PR 14-JUN-1996 US 08/663618
             PI PATRICK W GRAY
             PC C12N15/56,C12N9/24,C12N15/85,C12N5/10,C12N5/18,C07K16/40, PC
             A61K38/47
             CC Strandedness: Single;
             CC Topology: Linear;
             FH Key Location/Qualifiers
FEATURES    source
             1..30
             /organism="synthetic construct"
             /mol_type="genomic DNA"
             /db_xref="taxon:32630"
BASE COUNT  7 a 8 c 8 g 7 t
ORIGIN

Query Match      0.9%; Score 21; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2325 AGCTTGCATGCTGCAGGTGC 2345
Db  29 AGCTTGCATGCTGCAGGTGC 9

RESULT 10
BD085067/c
LOCUS       BD085067             30 bp      DNA             linear      PAT 27-AUG-2002
DEFINITION Synthesis of oligonucleotides.
ACCESSION  BD085067
VERSION    BD085067.1  GI:22630677
KEYWORDS   JP 2001525821-A/1.
SOURCE     unidentified
ORGANISM   unidentified
REFERENCE   1 (bases 1 to 30)
AUTHORS     Kwiatkowski,M., Landegren,U. and Nilsson,M.
TITLE       Synthesis of oligonucleotides
JOURNAL     Patent: JP 2001525821-A 1 11-DEC-2001;
COMMENT     MAREK KWIATKOWSKI
             OS Unidentified
             PN JP 2001525821-A/1
             PD 11-DEC-2001
             PF 14-MAY-1998 JP 1998549151
             PR 14-MAY-1997 SE 9701783-4
             PI MAREK KWIATKOWSKI,ULF LANDEGREN,MATS NILSSON

```

PC C07H21/00, C07H21/02, C07H21/04
CC Synthesis of oligonucleotides
FH Key Location/Qualifiers
FT source 1..30
/organism="Unidentified".
/Location/Qualifiers
1..30
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644" 6 t
BASE COUNT 6 a 9 c 9 g
ORIGIN
Query Match 0.9%; Score 21; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2325 AGCTTGCATGCTGCAGGTGCG 2345
|||||
Db 23 AGCTTGCATGCTGCAGGTGCG 3
RESULT 11
BD135123/c
LOCUS BD135123 30 bp DNA linear PAT 18-SEP-2002
DEFINITION Chitin-binding fragment of chitinase.
ACCESSION BD135123
VERSION BD135123.1 GI:23230068
KEYWORDS JP 2002505882-A/9.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 30)
AUTHORS Gray, P.W. and Tjoelker, L.W.
TITLE Chitin-binding fragment of chitinase
JOURNAL Patent: JP 2002505882-A 9 26-FEB-2002;
ICOS CORP.
OS Artificial Sequence
PN JP 2002505882-A/9
PD 26-FEB-2002
PR 12-MAR-1999 JP 2000535757
PI PATRICK W GRAY, LARRY W TJOELKER
PC C12N15/09, A61K38/00, A61K39/395, A61K45/00, A61P31/10,
PC C07K16/40,
PC C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N9/24, C12P21/08, G01N33/
PC 577,
PC C12N15/00, A61K37/02, A61K37/54, C12N5/00
CC Description of Artificial Sequence: primer
FH Key Location/Qualifiers
FT source 1..30
/organism="Artificial Sequence".
/Location/Qualifiers
1..30
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630" 7 t
BASE COUNT 7 a 8 c 8 g
ORIGIN
Query Match 0.9%; Score 21; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2325 AGCTTGCATGCTGCAGGTGCG 2345
|||||
Db 29 AGCTTGCATGCTGCAGGTGCG 9
RESULT 12
BD178195
LOCUS BD178195 30 bp DNA linear PAT 16-APR-2003
DEFINITION Novel DNA polymerase accessory factor.

ACCESSION BD178195
VERSION BD178195.1 GI:30015459
KEYWORDS JP 2002315588-A/8.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 30)
AUTHORS Ishino, Y. and Daimon, K.
TITLE Novel DNA polymerase accessory factor
JOURNAL Patent: JP 2002315588-A 8 29-OCT-2002;
BIOMOLECULAR ENGINEERING RESEARCH INSTITUTE
COMMENT OS Artificial Sequence
PN JP 2002315588-A/8
PD 29-OCT-2002
PF 02-NOV-2001 JP 2001337530
PI YOSHIZUMI ISHINO, KATSUYA DAIMON
PC C12N15/09, C07K14/195, C12N15/00
CC Novel DNA polymerase accessory factor
FH Key Location/Qualifiers
FT source 1..30
/organism="Artificial Sequence".
/Location/Qualifiers
1..30
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630" 7 t
BASE COUNT 5 a 10 c 8 g
ORIGIN
Query Match 0.9%; Score 21; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2325 AGCTTGCATGCTGCAGGTGCG 2345
|||||
Db 5 AGCTTGCATGCTGCAGGTGCG 25
RESULT 13
A23170
LOCUS A23170 32 bp mRNA linear PAT 30-NOV-1994
DEFINITION Artificial construct primer.
ACCESSION A23170
VERSION A23170.1 GI:641658
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 32)
AUTHORS Kaluza, B., Betzi, G. and Weidle, U.H.
TITLE Antibody expression vector and antibodies produced thereby
JOURNAL Patent: EP 0547631-A 7 23-JUN-1993;
BOEHRINGER MANNHEIM GMBH
FEATURES
source 1..32
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630" 8 t
BASE COUNT 6 a 9 c 9 g
ORIGIN
Query Match 0.9%; Score 21; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2325 AGCTTGCATGCTGCAGGTGCG 2345
|||||
Db 6 AGCTTGCATGCTGCAGGTGCG 26
RESULT 14
AR126240/c
LOCUS AR126240 32 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 2 from patent US 6180111.
ACCESSION AR126240
VERSION AR126240.1 GI:14112833
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 32)
AUTHORS Stein,D.C. and Stover,C.K.
TITLE Vaccine delivery system
JOURNAL Patent: US 6180111-A 2 30-JAN-2001;
FEATURES Location/Qualifiers
source 1..32
/organism="unknown"
BASE COUNT 7 a 8 c 10 g 7 t
ORIGIN

Query Match 0.9%; Score 21; DB 6; Length 32;
Best Local Similarity 100.0%; Pred.No. 5.3; 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;
Qy 2325 AGCTTGATGCTGCAGGTCG 2345
|||||
Db 29 AGCTTGATGCTGCAGGTCG 9

RESULT 15
AR264638
LOCUS AR264638 36 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 4 from patent US 6491908.
ACCESSION AR264638
VERSION AR264638.1 GI:29692909
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 36)
AUTHORS Rosenberg,A.S.
TITLE Selective elimination of T cells that recognize specific
preslected targets
JOURNAL Patent: US 6491908-A 4 10-DEC-2002;
FEATURES Location/Qualifiers
source 1..36
/organism="unknown"
BASE COUNT 8 a 10 c 9 g 9 t
ORIGIN

Query Match 0.9%; Score 21; DB 6; Length 36;
Best Local Similarity 100.0%; Pred.No. 5.4; 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;
Qy 2325 AGCTTGATGCTGCAGGTCG 2345
|||||
Db 14 AGCTTGATGCTGCAGGTCG 34

Search completed: November 14, 2003, 16:27:33
Job time : 8574.5 secs

THIS PAGE BLANK (USPTO)